

Pro-active Reservation Mechanisms for Next Generation Optical Networks

Huifang KONG

SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Supervised by Dr. Chris Phillips

Department of Electronic Engineering

Queen Mary, University of London

February 2006

To Chengmiao

Abstract

Next generation optical network design is driven by the rapid growth of Internet data traffic, and is influenced by the emergence of various services that have different requirements on delay tolerance, loss, and bandwidth. Optical Burst Switching (OBS) and Wavelength-Routed OBS (WR-OBS) are two possible solutions. However, OBS suffers from a relatively high blocking probability, resulting in the triggering of Transmission Control Protocol (TCP) retransmission timeouts and unpredictable delays. WR-OBS impedes the support for delay-sensitive service provisioning due to the inefficiency of the reservation protocol.

Researches have shown that the pro-active resource reservation schemes can achieve latency reduction and QoS provisioning in classical OBS networks. In this thesis, an in-depth study is carried out to explore the effects of various pro-active reservation mechanisms on next generation optical networks.

This work firstly proposes a pre-booking mechanism for the classical OBS networks. This work then extends existing research from small-timescale burst-level reservations to the large-timescale “average load” reservations in a hybrid network architecture, and quantifies their benefits and limitations. The proposed reservation mechanism is further investigated for use in connectivity acknowledged OBS networks, such as the WR-OBS networks, providing a significant enhancement to reactive two-way reservation protocols. Combined with the pro-active reservation concept, self-similar traffic statistics and traffic prediction technologies are studied.

In addition, with reference to optimal lightpath placement, this study examines existing Routing and Wavelength Assignment (RWA) algorithms, and proposes a two-stage RWA mechanism to obtain the optimal or near optimal solution within short timescales. The findings of this research should be directly relevant to the optimization of future optical networks.

Acknowledgement

My foremost gratitude belongs to my supervisor, Dr. Chris Phillips, who has not only provided academic supervision, support, and persistent encouragement, but also financially funded my PhD during the last three years, at Queen Mary, University of London.

I would also like to thank Dr. Raul Mondragon, Dr. John Schormans, and Professor Jonathan Pitts for their academic discussion and help.

I am deeply grateful to Robert Friskney at Nortel Networks, Ltd, and Song Dong who previously pursued a PhD in this department and currently works in France Telecom, UK. They have offered me lots of technical discussion, knowledge, and advice.

I would also like to thank the support staff in the Electronic Engineering department, especially Lynda Rolfe, Michele Pringle, Kok Ho Huen, George Cunliffe and Phil Willson, as well as other members of the team for their help with administrative work and computer services.

I would like to extend my acknowledgement to my friends in the department and outside the university. Thanks to Karen Shoop, Vindya Amaradasa, Constantinos Neophytou, Ning Lu, Xuefei Li, Landong Zuo, Xuan Huang, Huai Huang, Xiaonan Yang, Ling Lu, Rupert Ogilvie, etc. My study would not have been so pleasant without their friendship. I'd also like to thank Clare Dorothy for giving me the opportunity of teaching her some Chinese, which in turn improved my English.

Finally, with my love and gratitude, I would like to dedicate this thesis to my family for their continuous support and encouragement. Special thanks go to my husband, Chengmiao CAI, who has shared every moment of happiness and sadness with me during the last three years.

Table of Contents

ABSTRACT	3
ACKNOWLEDGEMENT	4
TABLE OF CONTENTS	5
LIST OF FIGURES	10
LIST OF TABLES	13
GLOSSARY	14
CHAPTER 1 INTRODUCTION	18
1.1 PROBLEM STATEMENT	18
1.2 RESEARCH CONTRIBUTIONS	19
1.3 ORGANISATION OF THE THESIS	20
CHAPTER 2 INTRODUCTION OF OPTICAL NETWORKS	22
2.1 KEY TECHNOLOGIES IN OPTICAL NETWORKS.....	22
2.1.1 WDM/DWDM.....	22
2.1.2 SONET/SDH.....	23
2.1.3 All-Optical Networking	25
2.2 WAVELENGTH ROUTED OPTICAL NETWORKS (WRON)	28
2.2.1 Features of WRON	29
2.2.2 Routing and Wavelength Assignment (RWA) in WRON.....	30
2.3 OPTICAL PACKET SWITCHING (OPS)	33
2.3.1 OPS Operations.....	34
2.3.2 Hardware Challenges for OPS.....	35
2.4 OPTICAL BURST SWITCHING (OBS)	38
2.4.1 OBS Operations.....	39
2.4.2 Burst Assembly	42
2.4.3 OBS Signalling Protocols.....	43
2.4.4 Routing and Wavelength Assignment (RWA) in OBS.....	47
2.4.5 Congestion Resolutions in OBS.....	48
2.4.6 QoS Provisioning in OBS.....	49
2.5 WAVELENGTH-ROUTED OBS	49

2.5.1	<i>Architecture of WR-OBS Networks</i>	50
2.5.2	<i>Burst Aggregation Mechanisms in WR-OBS</i>	51
2.5.3	<i>QoS Provisioning in WR-OBS</i>	51
2.5.4	<i>Provision of End-to-End Delay Guarantees in WR-OBS Networks</i>	53
2.6	SUMMARY	54
CHAPTER 3	EXISTING WORK ON PREDICTIVE RESERVATION MECHANISMS.....	55
3.1	FORWARD RESOURCE RESERVATION (FRR) MECHANISM.....	55
3.1.1	<i>Aggressive Resource Reservation</i>	56
3.1.2	<i>FRR for QoS Provisioning</i>	57
3.1.3	<i>Linear Predictive Filter (LPF) for Traffic Prediction</i>	58
3.2	CHALLENGES AND RESEARCH FOCUSES OF THIS THESIS	61
CHAPTER 4	PRE-BOOKING MECHANISM FOR CLASSICAL OBS NETWORKS.....	62
4.1	PROPOSED NETWORK ARCHITECTURE.....	62
4.1.1	<i>SP's Typical Operations</i>	63
4.1.2	<i>CNO's Operations</i>	64
4.2	SIMPLE PRE-BOOKING SCENARIO.....	68
4.2.1	<i>Simulation Model</i>	69
4.2.2	<i>Validation of Simulation Model</i>	72
4.2.3	<i>Performance Results and Evaluation</i>	73
4.3	SHIFTING MECHANISM EVALUATION.....	75
4.3.1	<i>Performance Results</i>	76
4.4	SUMMARY	77
CHAPTER 5	PRE-BOOKING MECHANISM FOR A HYBRID NETWORK ARCHITECTURE AND THE RWA MECHANISM.....	78
5.1	INTRODUCTION.....	78
5.2	PROPOSED NETWORK ARCHITECTURE.....	78
5.2.1	<i>Service Providers (SPs)</i>	79
5.2.2	<i>Carrier Network Operator (CNO)</i>	81
5.2.3	<i>Potential Benefits</i>	82
5.3	RWA ALGORITHMS	82
5.3.1	<i>Problems under Study</i>	83

5.3.2	<i>Conventional ILP Models</i>	83
5.3.3	<i>Proposed Approach for Wavelength Constrained Max-RWA Problem</i>	94
5.4	SIMULATION MODELS FOR PROPOSED NETWORK ARCHITECTURE	102
5.4.1	<i>Central Network Planning Node</i>	103
5.4.2	<i>Edge Nodes</i>	104
5.4.3	<i>Core Optical Switch</i>	105
5.4.4	<i>Validation of the Simulation Models</i>	105
5.5	SIMULATION RESULTS FOR PROPOSED NETWORK ARCHITECTURE	106
5.5.1	<i>Blocking Probability versus the Load</i>	106
5.5.2	<i>Blocking Probability versus the Allowed Extra Edge Delay</i>	110
5.6	NUMERICAL RESULTS FOR ILP-BASED RWA MODELS.....	110
5.6.1	<i>Results of Conventional ILP Models without the Wavelength Continuity Constraint</i>	111
5.6.2	<i>Performance for the Proposed Two-Stage Approach</i>	113
5.7	SUMMARY	115
CHAPTER 6 PRE-BOOKING MECHANISM FOR END-TO-END CONNECTIVITY		
ACKNOWLEDGED OBS NETWORKS.....		
6.1	INTRODUCTION.....	117
6.2	NETWORK ARCHITECTURE	118
6.3	EXPLICIT AND IMPLICIT PRE-BOOKING	120
6.4	BURST ASSEMBLY MECHANISMS IN IMPLICIT PRE-BOOKING.....	121
6.5	IMPLICIT PRE-BOOKING SIGNALLING WITH CENTRALIZED CONTROL.....	121
6.5.1	<i>Traditional Signalling and Reservation</i>	121
6.5.2	<i>Proposed Signalling and Reservation</i>	123
6.5.3	<i>RWA Algorithm in Pre-booking Mechanism</i>	124
6.5.4	<i>Comparison with WR-OBS</i>	126
6.6	IMPLICIT PRE-BOOKING SIGNALLING WITH DISTRIBUTED CONTROL.....	127
6.6.1	<i>Traditional Signalling and Reservation</i>	127
6.6.2	<i>Proposed Signalling and Reservation</i>	128
6.7	ANALYSIS OF LIGHTPATH BANDWIDTH EFFICIENCY	129
6.7.1	<i>Centralized Schemes</i>	129
6.7.2	<i>Distributed Schemes</i>	131

6.8 BURST LENGTH STUDY WITH SELF-SIMILAR INCOMING TRAFFIC	132
6.8.1 Definition of Self-Similarity.....	133
6.8.2 Scaling Behaviours in Internet Traffic.....	134
6.8.3 Burst Length Statistics.....	135
6.9 RESERVATION STRATEGY BASED ON BURST LENGTH PREDICTION.....	135
6.9.1 Burst Length Reservation Strategy with LBS Assembly Mechanism	136
6.9.2 Burst Length Reservation Strategy with UBS Assembly Mechanism.....	137
6.10 BIT LOSS RATE CAUSED BY INSUFFICIENT RESERVATION	137
6.11 SUMMARY	139
CHAPTER 7 SIMULATION MODELS AND RESULTS FOR PRE-BOOKING MECHANISM IN CONNECTIVITY ACKNOWLEDGED OBS NETWORKS	140
7.1 SIMULATION MODELS	140
7.1.1 H-Self-Similar Traffic Modelling.....	141
7.1.2 Central Node Model.....	144
7.1.3 Edge Node Model.....	148
7.1.4 Validation of the Simulation Models.....	152
7.2 SIMULATION RESULTS FOR SELF-SIMILAR TRAFFIC TRACES	153
7.2.1 Simulated Burst Length Distribution.....	153
7.2.2 Bit Loss Rate Caused by Insufficient Reservation	155
7.3 NETWORK SIMULATION ENVIRONMENT	157
7.4 NETWORK SIMULATION RESULTS WITH CBR TRAFFIC	158
7.4.1 Results for Scenario-1 with a Single RWA Attempt	159
7.4.2 Results for Scenario-2 with Multiple RWA Attempts	164
7.5 NETWORK SIMULATION RESULTS WITH SELF-SIMILAR TRAFFIC	171
7.5.1 Performances versus the Traffic Load.....	172
7.5.2 Performance versus the k Value for Pre-booking Mechanism.....	177
7.6 SUMMARY	179
CHAPTER 8 CONCLUSIONS.....	181
8.1 DISCUSSION.....	181
8.1.1 Pre-booking Mechanism for Classical OBS Networks	182
8.1.2 Pre-booking Mechanism for a Hybrid Network Architecture and RWA Mechanism.....	182

8.1.3 <i>Pre-booking Mechanism for Connectivity Acknowledged OBS Networks</i>	183
8.2 FUTURE WORK.....	183
APPENDIX A – AUTHOR’S PUBLICATIONS	187
APPENDIX B – NETWORK TOPOLOGIES	188
APPENDIX C – TRAFFIC MATRICES FOR RWA STUDY	190
APPENDIX D - MATHEMATICAL MODELS FOR STANDARD JET SYSTEM AND PRE-BOOKING MECHANISM	192
APPENDIX E – VALIDATION OF SIMULATION MODELS FOR PRE-BOOKING MECHANISM IN CLASSICAL OBS NETWORKS	193
APPENDIX F - VALIDATING THE SELF-SIMILAR TRAFFIC	196
REFERENCES	199

List of Figures

Fig 2-1: Wavelength Division Multiplexing (WDM)	22
Fig 2-2: A connection in the SONET/SDH ring	23
Fig 2-3: Channels in a fibre	24
Fig 2-4: A schematic structure of Optical Add-Drop Multiplexer (OADM).....	26
Fig 2-5: A schematic structure of Optical Cross-Connect (OXC).....	27
Fig 2-6: An example of wavelength routed optical network.....	29
Fig 2-7: A generic structure of the OPS OXC.....	35
Fig 2-8: OBS network structure.....	39
Fig 2-9: General structure of the OBS edge node.....	40
Fig 2-10: General structure of the OBS core node.....	41
Fig 2-11: Offset time in JET protocol	45
Fig 2-12: JIT signalling protocol	46
Fig 2-13: WR-OBS network architecture	50
Fig 2-14: Request server in WR-OBS control node.....	52
Fig 3-1: FRR mechanism.....	56
Fig 3-2: QoS provisioning in FRR mechanism.....	58
Fig 3-3: LMS-based predictor.....	60
Fig 4-1: Proposed Architecture for Classical OBS	62
Fig 4-2: An Example – Negotiation and Signalling.....	65
Fig 4-3: Cost Gradient on a Given Link	66
Fig 4-4: Example Resource Reservation Table in the Core Optical Switch	67
Fig 4-5: Edge Node Model	69
Fig 4-6: Handling the Predicted Traffic	70
Fig 4-7: Handling the non-predicted traffic	71
Fig 4-8: Core Node Model.....	71
Fig 4-9: Workflow in the Core Node Model.....	72
Fig 4-10: Simple Pre-booking Performance with a Constant Traffic Mixture.....	73
Fig 4-11: Simple Pre-booking Performance with Various Traffic Mixtures	74

Fig 4-12: Shift Scheme Performance with a Constant Traffic Mixture	76
Fig 4-13: Shift Scheme Performance with Various Traffic Mixtures.....	77
Fig 5-1: Proposed Architecture with Hybrid Operations	79
Fig 5-2: Functionality Blocks in a Service Provider	79
Fig 5-3: An Example of Resource Reservation Strategy	80
Fig 5-4: Comparison of Model Size between the Flow-based and the Source-based Models....	94
Fig 5-5: Demonstration of Graph Colouring.....	97
Fig 5-6: An example of BB search with chronological backtracking	99
Fig 5-7: An example of colouring graph.....	100
Fig 5-8: Network Structure of Proposed Hybrid Network Architecture	102
Fig 5-9: Node Model of the Central Network Planning Node	103
Fig 5-10: Blocking Probability versus the Load in a Network without Wavelength Conversion Support.....	107
Fig 5-11: Blocking Probability versus the Load in a Network with Wavelength Conversion Support for OBS Service	109
Fig 5-12: Blocking Probability versus the Allowed Extra Edge Delay.....	110
Fig 6-1: Proposed network architecture	119
Fig 6-2: Traditional lightpath signalling in centralized control plane.....	122
Fig 6-3: Pre-booking signalling time parameters in centralized control plane	123
Fig 6-4: Pre-booking Lightpath signalling in centralized control plane	124
Fig 6-5: RWA algorithm in pre-booking mechanism	125
Fig 6-6: Traditional LSP reservation in distributed control plane.....	127
Fig 6-7: Pre-booking LSP reservation in distributed control plane.....	128
Fig 6-8: Lightpath bandwidth efficiency in centralized control plane	130
Fig 6-9: Lightpath bandwidth efficiency in distributed control plane	132
Fig 7-1: The ON-OFF Process	142
Fig 7-2: Superposition of ON-OFF processes.....	143
Fig 7-3: Central Node Process Model for WR-OBS and Traditional Reservation Schemes	145
Fig 7-4: Central Node Process Model for Pre-booking Scheme	147
Fig 7-5: Node Model for edge node.....	148
Fig 7-6: Edge node process model for WR-OBS and traditional reservation schemes	149

Fig 7-7: Edge node process model for pre-booking scheme.....	151
Fig 7-8: PDF for aggregation burst length from 10 and 100 ON-OFF processes.....	154
Fig 7-9: Mean and variance of aggregated burst length from 10 ON-OFF processes	154
Fig 7-10: Bit loss rate from 10 ON-OFF processes	155
Fig 7-11: Bit loss rate from 100 ON-OFF processes	156
Fig 7-12: Burst Blocking Probability versus Traffic Load for Scenario-1.....	160
Fig 7-13: Lightpath Bandwidth Efficiency versus Traffic Load for Scenario-1	160
Fig 7-14: Burst Blocking Probability versus the Number of Wavelengths for Scenario-1	161
Fig 7-15: Burst Blocking Probability versus the Network Size for Scenario-1	163
Fig 7-16: Lightpath Bandwidth Efficiency versus the Network Size for Scenario-1	163
Fig 7-17: Burst Blocking Probability versus the Load for Scenario-2.....	166
Fig 7-18: Lightpath Bandwidth Efficiency versus the Load for Scenario-2.....	168
Fig 7-19: Burst Blocking Probability versus the Number of Wavelengths for Scenario-2	169
Fig 7-20: Burst Blocking Probability versus the Network Size for Scenario-2.....	170
Fig 7-21: Burst Blocking Probability versus the Traffic Load with Self-similar Traffic	172
Fig 7-22: Bit Blocking Probability versus the Traffic Load with Self-similar Traffic.....	174
Fig 7-23: Lightpath Bandwidth Efficiency versus the Traffic Load with Self-similar Traffic..	176
Fig 7-24: Burst Blocking Probability versus the k Value in Pre-booking Mechanism with Self-Similar Traffic	177
Fig 7-25: Bit Blocking Probability versus the k Value in Pre-booking Mechanism with Self-Similar Traffic	178

List of Tables

Table 5-1: Parameters for Path-based ILP Formulations	84
Table 5-2: Parameters for Flow-based ILP Formulations	86
Table 5-3: Number of Variables and Constraints for the Three Models	92
Table 5-4: Objective Value from the Wavelength unconstrained ILP Models	111
Table 5-5: Execution Time for the Wavelength unconstrained ILP Models	112
Table 5-6: Performance of Proposed RWA Mechanism.....	114

Glossary

ADM	Add Drop Multiplexer
AR	Auto-Regression
ARMA	Auto-Regression and Moving Average
ARIMA	Auto-Regression-Integrated Moving Average
ATM	Asynchronous Transfer Mode
AU	Administration Units
AUR-E	Adaptive Unconstrained Routing-Exhaustive algorithm
BB	Branch-and-Bound
BHP	Burst Header Packet
CAC	Connection Admission Control
CBR	Constant-Bit-Rate
CD	Chromatic Dispersion
CDF	Cumulative Distribution Function
CLT	Central Limit Theorem
CNO	Carrier Network Operator
CoS	Class of Service
CR-LDP	Constraint-based Routed LDP
DFS	Depth-First-Search
DR	Delayed Reservation
DWDM	Dense Wavelength Division Multiplexing
EDF	Earliest Deadline First
EDFA	Erbium Doped Fibre Amplifier
F-ARIMA	Fractal-ARMA
FBG	Fibre Bragg Grating
FDL	Fibre Delay Line
FEC	Forwarding Equivalence Class
FFT	Fast Fourier Transform
FGN	Fractional Gaussian Noise
FRR	Forward Resource Reservation

GMPLS	Generalized MPLS
HDLC	High-level Data Link Control
ILP	Integer Linear Programming
ISO	International Standards Organisation
ISP	Internet Service Provider
ITU	International Telecommunication Union
JET	Just-Enough-Time
JIT	Just-In-Time
LAN	Local Area Network
LAUC-VF	Latest Available Unused Channel with Void Filling
LBS	Limited-Burst-Size
LDP	Label Distribution Protocol
LIB	Label Information Base
LMMSE	Linear Minimum Mean Square Error
LMS	Least Mean Square
LOBS	Label-switched OBS
LPF	Linear Predictive Filter
LRD	Long-Range-Dependence
LSP	Label Switched Path
LSR	Label Switched Router
MAPOS	Multiple Access Protocol over SONET/SDH
MEMS	Micro-Electromechanical Systems
MMSE	Minimum Mean Square Error
MPLS	Multi-Protocol Label Switching
NNI	Network-to-Network Interface
NP-hard	Non-Polynomially hard
OADM	Optical Add-Drop Multiplexer
OAM&P	network Operation, Administration, Maintenance, and Provisioning
OBS	Optical Burst Switching
OC	Optical Carrier
OEO	Optical-to-Electrical-to-Optical

OPS	Optical Packet Switching
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
OXC	Optical Cross-Connect
PDF	Probability Density Function
PMD	Polarisation Mode Dispersion
PPP	Point-to-Point Protocol
QoS	Quality of Service
RSVP	ReSerVation Protocol
RWA	Routing and Wavelength Assignment
SCU	Switch Control Unit
SDH	Synchronous Digital Hierarchy
SLA	Service Level Agreement
SLALOM	Semiconductor Laser Amplifier in Loop Optical Mirror
SMF	Single Mode Fibre
SOA	Semiconductor Optical Amplifier
SOH	Section Overhead
SONET	Synchronous Optical Network
SP	Service Provider
STM	Synchronous Transport Module
STS	Synchronous Transport Signal
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TOAD	Terahertz Optical Asymmetric Demultiplexer
TU	Tributary Units
TUG	Tributary Units Group
UBS	Unlimited-Burst-Size
VC	Virtual Container (in SDH)
VCC	Virtual Circuit Connection
VOA	Variable Optical Attenuations
VPC	Virtual Path Connection

WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
WR	Wavelength Routing
WRON	Wavelength Routed Optical Network
WR-OBS	Wavelength-Routed Optical Burst Switching
WRR	Weighted Round-Robin algorithm
WSS	Wavelength Selective Switch

Chapter 1 Introduction

1.1 Problem Statement

During the last ten years, the composition of the total network traffic has changed to a different picture [CO02][ESG01]. Data collected from US Internet backbones for last decade shows that the Internet data traffic has been doubling annually since 1997, whilst, traditional voice traffic has grown little (about 10% per year) for many decades [CO02]. This phenomenon has stimulated the design of various future optical network architectures, such as to enable the transition from voice-centric to converged networks supporting both the voice and the data traffic, as well as to improve the network resource utilization efficiency [ESG01] [CQY04][DB02][GRG98][XGY01][CAA01].

Future network architecture design is also influenced by the emergence of various services, which can have different requirements on tolerable delay, packet loss rate per application, and bandwidth [ESG01]. In addition, in combination with the packet loss, the latency plays an important role for provisioning of upper-layer applications such as TCP, whose throughput depends on both delay and packet loss. Therefore, it is very important to ensure the performance of currently proposed architectures in terms of loss, latency, and QoS provisioning, is sufficient.

Optical Burst Switching (OBS) architectures attempt to map the IP paradigm into optical core networks with reduced processing required for packet forwarding, and to achieve high multiplexing gain and greater flexibility in response to the changing traffic load [CQY04]. However, most of the existing OBS protocols attempt to transport data without the knowledge of end-to-end connectivity, thus suffering from high blocking probability, leading to Transmission Control Protocol (TCP) retransmission events, causing increased delays and congestion [SLK05].

Alternatively, Wavelength-Routed OBS (WR-OBS) combines the concept of OBS with the dynamic lightpath assignment under the fast circuit-switching technology. It offers many desirable features, such as the dynamic lightpath provisioning, the admission control, and the end-to-end guaranteed services [DB02][DB022][DZB04][MKB02]. However, in geographically large backbones, the inefficiency of the two-way reservation protocol is significant, and the

mechanism impedes latency sensitive service provisioning [MKB02].

Recently, a pro-active resource reservation scheme has been proposed to achieve the latency reduction and QoS differentiation in classical OBS¹ networks [LAO03]. In this thesis, an in-depth study is carried out to explore the effects of various pro-active reservation mechanisms on the next generation optical networks.

More specifically, in the classical OBS networks, this work firstly introduces a pre-booking mechanism with various forms of intelligence at the network edge. This work then extends the existing study from the small timescale burst level reservation to the large timescale average load level reservation in a hybrid network architecture, and quantifies the benefits and limitations of pro-active reservation mechanisms. In the connectivity acknowledged OBS networks, such as the WR-OBS networks, this work investigates the application of the proposed pre-booking mechanism, which provides an enhancement to the reactive two-way reservation protocols. Coupled with the pro-active reservation mechanisms, traffic prediction technology and self-similar traffic statistics are studied and analysed. In addition, with respect to optimal lightpath placement, this study looks into existing Routing and Wavelength Assignment (RWA) algorithms, and proposes a two-stage RWA mechanism to obtain the optimal or near optimal solution within a short timeframe. The findings of this research can help the optimisation of the future optical networks.

1.2 Research Contributions

The following contributions have been made by this study.

- ◆ A novel pre-booking mechanism is proposed for classical OBS networks. The mechanism suggests several forms of intelligence for both the Service Providers (SPs) and the Carrier Network Operator (CNO) to improve network performance, such as the overall blocking probability, the QoS provisioning, etc.
- ◆ The pre-booking mechanism is adapted and applied in a hybrid network architecture, where

¹ In this thesis, the classical OBS refers to the connectivity un-acknowledged OBS, such as the OBS based Joust Enough Time (JET) or Just In Time (JIT) protocol.

both the wavelength routing and the classical OBS are supported. The results indicate that the proposition achieves performance gain over the classical OBS in a network without wavelength conversion support. However, in the case when wavelength conversion becomes widely available, classical OBS will be a good option.

- ◆ To achieve the optimal lightpath placement, a two-stage Routing and Wavelength Assignment (RWA) infrastructure is proposed for wavelength constrained Max-RWA problem. The proposition overcomes the disadvantages of the direct Integer Linear Programming (ILP) models in terms of the problem solving time.
- ◆ Pre-booking for connectivity acknowledged OBS networks is studied. The mechanism is able to achieve a high lightpath bandwidth efficiency, which further results in a lower blocking probability and higher supported traffic loads for latency-sensitive traffic.
- ◆ The burst length from the self-similar incoming traffic and the corresponding reservation strategy are investigated. The bit loss rate associated with the reservation strategy is mathematically analysed and simulated.

1.3 Organisation of the Thesis

Chapter 2 provides background relating to the key technologies in optical networks. After the introduction of current optical network and future all-optical network, this chapter describes four potential switching technologies for the next generation optical networks. These switching technologies are the Wavelength Routed Optical Network (WRON), the Optical Packet Switching (OPS), the Optical Burst Switching (OBS), and the Wavelength-Routed OBS (WR-OBS), respectively.

Chapter 3 describes the existing work for the predictive resources reservation in classical OBS networks. This chapter then outlines the challenges and focuses for this thesis.

Chapter 4 proposes a pre-booking mechanism for classical OBS networks. The proposed network architecture adopts various forms of intelligence at the network edge, such as to improve the network performance, like the blocking probability, QoS differentiation, etc. Following the introduction of network architecture, the simulation models and results are

presented and evaluated in multiple scenarios.

Chapter 5 adapts the pre-booking mechanism into newly proposed hybrid network architecture, following which the RWA algorithms based on the ILP approach are investigated. In the second part of the chapter, the simulation models and results for the proposed architecture are evaluated in both the wavelength conversion supported and unsupported scenarios. In the end, the numerical results for the RWA approach are also presented.

Chapter 6 gives a description and an analysis for the pre-booking mechanism in connectivity acknowledged OBS networks. The signalling protocols and the corresponding lightpath bandwidth efficiency are described and compared in both the centralized and the distributed control planes. Additionally, this chapter also studies the burst length statistics with self-similar incoming traffic, and proposes a reservation strategy to reduce the bit loss caused by prediction error.

Chapter 7 presents the simulation results for self-similar models, and validates the analytical bit loss rate as described in chapter 6. This chapter also delineates the network simulation models and evaluates network performance of the pre-booking mechanism in connectivity acknowledged OBS networks. The mechanism is studied via various scenarios with both the Constant Bit Rate (CBR) traffic and the self-similar traffic.

Chapter 8 discusses the work in this thesis, draws conclusions, and considers future work.

Appendix A lists Author's publications associated with this study.

Appendix B provides the network topologies used in the simulation and other experiments.

Appendix C gives out the traffic matrices used in the RWA study.

Appendix D presents the mathematical models for classical OBS and the simple pre-booking mechanism.

Appendix E describes and presents validation results for the pre-booking models in classical OBS networks.

Appendix F examines the traffic traces, which are produced by self-similar traffic generation models.

Chapter 2 Introduction of Optical Networks

This chapter firstly delineates key technologies in both the current optical networks and the future All-Optical-Networks (AON). It then describes and compares optical switching technologies, including the Wavelength Routing (WR), the Optical Packet Switching (OPS), the Optical Burst Switching (OBS), and the Wavelength-Routed OBS (WR-OBS), in detail.

2.1 Key Technologies in Optical Networks

2.1.1 WDM/DWDM

WDM is a multiplexing method allowing many independent signals to be transmitted simultaneously on one fibre, but with each signal located at a different wavelength [Taga96] [Cav00]. Fig 2-1 schematically shows the WDM functions. Four wavelengths are emitted from independent lasers, and enter a WDM multiplexer. The separated signals are combined into one wavelength-modulated light beam, and are transmitted on the fibre. At the other end of the fibre, the combined signals are separated by a de-multiplexer, and received by the corresponding photo-detector.

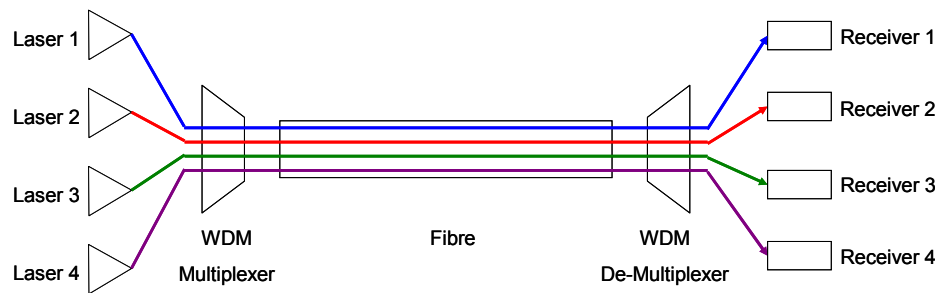


Fig 2-1: Wavelength Division Multiplexing (WDM)

The benefit of WDM/DWDM techniques is that they support multiple channels within one fibre, and increases the optical network capacity without manually laying more fibres. For example, the Alcatel 1626 Light Manager (LM) platform offers a capacity of up to 192 wavelength channels at 10 Gbps in the extended C-band², which achieves a total capacity up to 1.92Tbps

² Fibre optics uses the specific regions on the optical spectrum called windows. The most commonly used windows are the 1310 nm window called S band, the 1550 nm window called C band, and the 1625 nm window called L band. These three windows provide lower attenuations than other spectrums.

[Alc1626].

In high-capacity environments, the extension of WDM is the so-called Dense WDM (DWDM). DWDM places multiple wavelengths densely, and supports a large number of wavelengths in a single fibre. However, DWDM is sensitive to signal crosstalk, Chromatic Dispersion (CD), and Polarisation Mode Dispersion (PMD), etc.³, so it requires high-performance optical components, and makes the transmission system expensive [ZTZ05][Taga96]. Therefore, DWDM is usually deployed in the core/backbone network environment.

2.1.2 SONET/SDH

Time Division Multiplex-based Synchronous Optical NETwork (SONET)/the Synchronous Digital Hierarchy (SDH) protocol has become the de facto standard for backbones [Cav00].

2.1.2.1 SONET/SDH Rings

Modern SONET/SDH fibre optic transport networks are primarily constructed in rings with SONET/SDH terminals [Cav00].

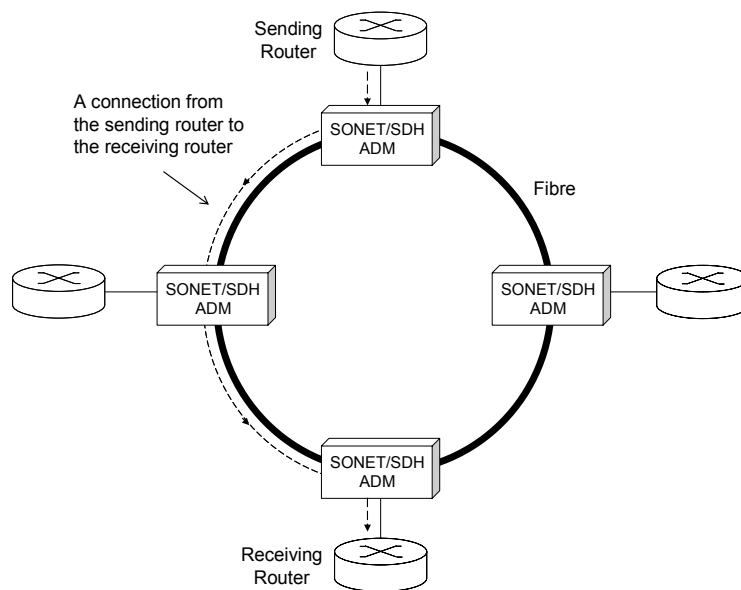


Fig 2-2: A connection in the SONET/SDH ring

Before the Wavelength Division Multiplexing (WDM) technology was invented, a fibre only

³ Crosstalk measures the amount of signal crossing between wavelength channels. Chromatic Dispersion is caused by different propagation speed of various wavelengths transmitting on the same fibre. PMD is the pulse disorder that is mainly affected by the asymmetry of the fibre optic strand.

supports one wavelength containing multiple time-slot channels. Establishing a connection means the provisioning of a point-to-point circuit between the devices that need to communicate. As shown in Fig 2-2, when the traffic enters the SONET/SDH ring, the transmitting router sends traffic to a SONET/SDH ADM, where the ADM inserts the traffic into a TDM time slot on the ring. When the traffic arrives at the destination, the receiving ADM removes the traffic from the ring and forwards it to the receiving router.

With WDM in SONET/SDH rings, the connections are similar to those in single-wavelength-based SONET/SDH rings [Cav00]. As shown in Fig 2-3, the fibre with a single wavelength consists of multiple TDM channels. The fibre with multiple wavelengths resembles the aggregation of the fibres with a single wavelength, and each wavelength channel can be further multiplexed into separate TDM channels. Therefore, the operations of the WDM-based SONET/SDH rings are similar as the ones illustrated in Fig 2-2, but the system requires extra optical components, such as the Optical Add/Drop Multiplexer (OADM), to either add or drop a wavelength, and the optical combiner/splitter to multiplex/demultiplex wavelengths.

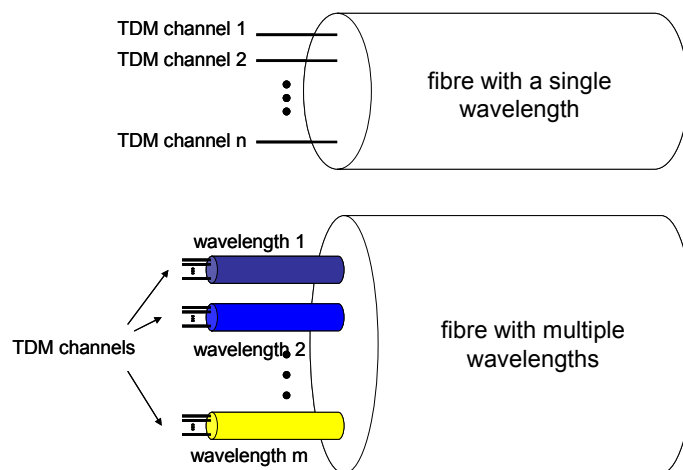


Fig 2-3: Channels in a fibre

2.1.2.2 The Static Feature of the SONET/SDH Rings

In legacy SONET/SDH rings that do not employ WDM technologies, connections between end-nodes are usually maintained for at least one month⁴. However, the operations for

⁴ The life time of a connection is affected by the policies of network operator, like the British Telecommunications

connection setup and teardown can be automatically controlled via a connection management platform [CJ05].

In SONET/SDH rings with WDM technology, the connection establishment and release can hardly be automatic. This is mainly due to the physical layer constraints including the following facts.

Firstly, most of the commercial optical system components in use, such as the transmitters, multiplexers, amplifiers, filters (OADM), and receivers, are operating with fixed wavelengths [TZT03]. These equipments have no flexibility in wavelength selection or wavelength allocation.

Secondly, to achieve reasonable reach, the powers of each wavelength must be controlled throughout the system. Given that amplifiers do not provide uniform gain across the spectrum, this requires frequent power re-balancing. Especially in OADM, when the wavelengths are dropped or added, the power of the wavelength channels will be altered, so Variable Optical Attenuations (VOA) are used to adjust the power level of those channels [Rob01]. Currently, VOA are mostly based on manual configuration using a screwdriver, so reconfigurable OADMs are not commercially employed yet.

2.1.3 All-Optical Networking

Currently, the operation of commercial optical networks still involves electronic processing, which causes a mismatch between the electronic processing speeds and the optical transmission rates, and increases the capital and operational costs. All-Optical Networking aims to dismiss the electronic processing in the future optical networks, and also provide flexibility and efficiency from traffic engineering point of view. More specifically, future optical networks should not only operate on rings or lines, but also on any mesh topology. Future optical network architecture should also appropriately accommodate the Internet data traffic, and at the same time provide QoS support to voice, video, and other types of service. Additionally, as Internet data traffic incoming bit rate can fluctuate during a short period of time [EA97], it is very important to have a dynamic architecture that can quickly react to the changing traffic load and

ple, who normally leases a connection to a customer for at least one month.

improve the efficiency of resource utilization.

The work on all-optical networking progresses gradually, and the effort mainly concentrates on the physical optical component development and on the network architecture design.

2.1.3.1 Optical Components

All-optical networking requires reliable, high-speed optical equipments. Apart from the laser, and optical amplifier, Optical Add-Drop Multiplexer (OADM) and Optical Cross-Connect (OXC) play an important role in flexible lightpath⁵ provisioning and realising the transparent optical networks [GRS00].

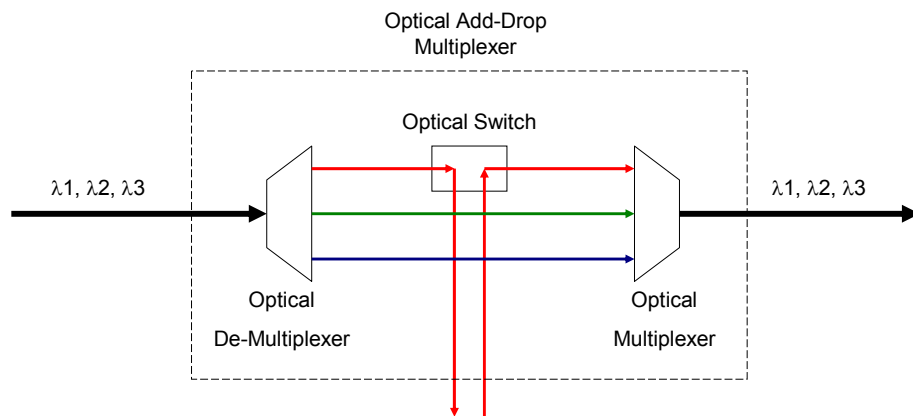


Fig 2-4: A schematic structure of Optical Add-Drop Multiplexer (OADM)

OADM is designed to insert, remove, or pass wavelength channels in the WDM/DWDM transmission system, without undergoing the signal conversion from optical to electrical domain. Fig 2-4 provides a schematic presentation of the OADM, where a wavelength is extracted or inserted by an optical switch entity⁶, whilst other wavelengths are intact. The early emerged OADMs are mostly realized by the Fibre Bragg Grating (FBG) filters, but they are confined to fixed add/drop or through wavelength channels [TZT03][DJP02]. Due to the technical evolution in Wavelength Selective Switch (WSS) and tuneable filters, reconfigurable OADMs come to

⁵ A lightpath is typically a high-bandwidth connection carrying data at up to several Gigabits per second.

⁶ The optical switch entity in the OADM can have various forms. For example, they can be a Micro-Electro-Mechanical Systems (MEMS)-based mirror substrate, or optical filters, depending on the design and application demands.

exist and provide more flexibility in lightpath provisioning [APP01][MNG05].

Optical Cross-Connect (OXC) is a more powerful and complicated optical switching equipment than the OADM. As shown in Fig 2-5, OXC cross-connects a number of fibre pairs via the optical switch plane, enabling routing of any incoming wavelength channels to the appropriate output ports. The optical switch plane can be realized by many technologies, such as the FBG-based blocks [CL98], whilst, the most promising technologies are the Micro-Electro-Mechanical Systems (MEMS)-based mirror substrate [YYT05]. The OXC has the ability to route all the wavelengths on an incoming fibre to a different outgoing fibre. It can also selectively direct a specific wavelength from an incoming fibre to another. If the OXC is coupled with wavelength converters, it is able to switch one wavelength to another wavelength. Meanwhile, most of the OXC design also supports local wavelength add/drop functionalities.

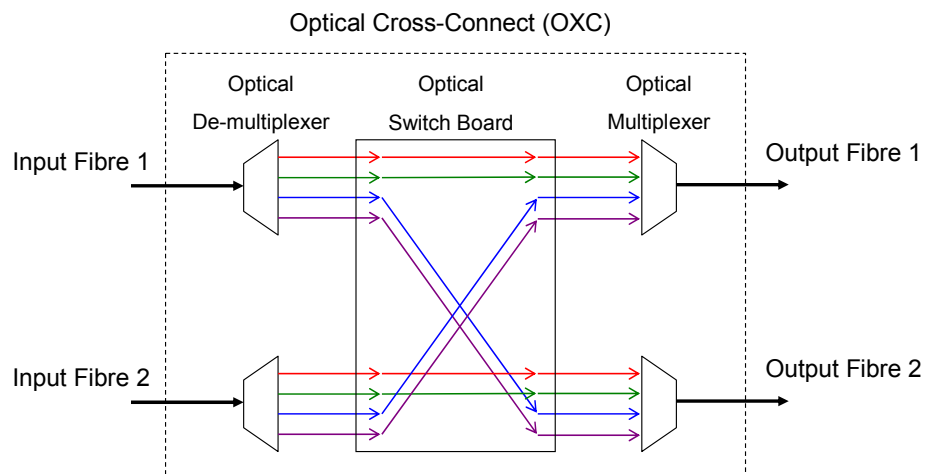


Fig 2-5: A schematic structure of Optical Cross-Connect (OXC)

Although more and more advanced optical devices have been developed, the current-state-of-the-art is still insufficient for realizing all-optical networks. For example, large scale OXCs and OADMs may suffer from high insertion loss⁷ and crosstalk. Fully integrated optical processing and optical wavelength conversion is very hard to achieve [GSP00][Yoo96]. Additionally, the reported switching speed to date is still in millisecond region, hardly satisfying the requirements of optical packet switching [RKH01][YLG01].

⁷ Insertion loss is the signal attenuation caused by the components inside the component.

2.1.3.2 Next Generation Optical Network Architectures

All-optical networking not only involves hardware developments, but also advances in network architecture design.

So far, Wavelength-Routed Optical Network (WRON), Optical Packet Switching (OPS), Optical Burst Switching (OBS), and Wavelength Routed OBS (WROBS) are the potential candidates for the next generation optical networks. Each of them operates at a different time granularity, thus putting different demands on the optical devices in terms of the switching speed, optical buffering and processing capacity, etc.

Briefly, WRON assumes that traffic pattern is rather static, so it rarely reconfigures the optical switches, putting low requirements on the switching speed. However, WRON is relatively insensitive to changing traffic load.

OPS aims to forward traffic on a packet-by-packet basis, so it requires high switching performance and considerable optical buffering. Assuming that the average packet length is 389.5Bytes [XVC00], and the channel transmission rate is 10Gbps, then about 3.2 million packets have to be switched each second, requiring a switching speed of about 311ns per packet. If the switching speed is too low, large capacity optical buffers, typically realised as Fibre Delay Lines, are required. Therefore, OPS poses a considerable challenge for optical device manufacturers.

OBS and WR-OBS attempt to forward packets with reduced header processing. They assemble multiple packets into bigger bursts, and use a separate signalling header to reserve resources ahead of each data burst. As the transmission granularity in OBS and WR-OBS is larger than the OPS, the required switching speed is reduced, thus increasing the technologies' feasibility. More details of the four switching technologies are described in the following sections.

2.2 Wavelength Routed Optical Networks (WRON)

WRON is the network architecture based on the wavelength routing technology. Wavelength routing is an optical version of circuit switching, where a dedicated end-to-end lightpath is established between specified ingress and egress nodes. More specifically, in circuit switching, an end-to-end connection is committed as long as at least one circuit is available on each link

along the selected path. In wavelength routing, wavelengths are the counterpart of the circuits as in the circuit switching [Sue05].

In reality, wavelength converters, enabling wavelength conversion, are very expensive, so the established end-to-end lightpath is usually assigned the same unique wavelength throughout of the whole span; this is called the wavelength continuous lightpath.

2.2.1 Features of WRON

Fig 2-6 shows an example of the WRON. The network is connected by OXCs with fibres containing multiple wavelengths. Three end-to-end lightpaths are established from OXC1 to OXC3, OXC1 to OXC5, and OXC4 to OXC6, respectively. Each lightpath occupies the same wavelength at each link along the path.

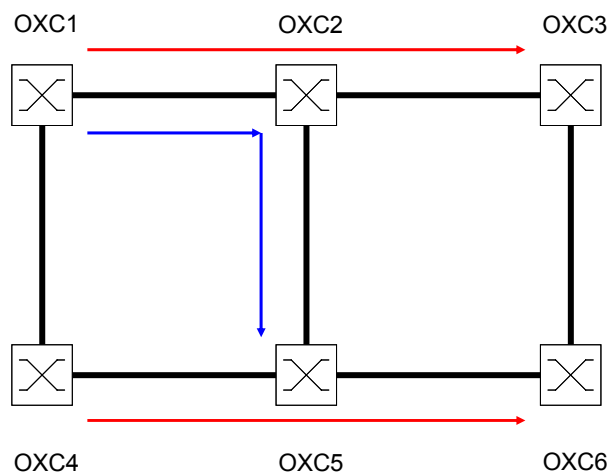


Fig 2-6: An example of wavelength routed optical network

The lightpath is dedicatedly allocated for the specified end-nodes. Once entering the first OXC along the lightpath, traffic will be transparently delivered to the other end. The traffic will not be terminated or inserted at any intermediate nodes along the lightpath. Furthermore, there are no buffers in the core nodes, and buffers are only available at the network edge in the electronic form. In other words, when a packet arrives at the network edge, it might be buffered for a while waiting for a free lightpath for the required source-destination. Therefore, traffic loss mainly occurs at the network edge nodes, and the end-to-end delay mainly includes the edge buffering delay and the propagation delay.

Most of the WRON assumes that traffic load for the whole network is known *a priori*, and, once

the lightpaths are established, they will remain in place for months or even years. As reconfiguration rarely takes place, there are little constraints on OXC's switching speed, and WRON is relatively easy to design and operate. However, WRON is not sensitive to varying traffic loads, causing resource over-provisioning when the load is low, and resource starvation when the load is high. Additionally, WRON provides the whole wavelength bandwidth to a lightpath (e.g. 10 or 40Gbps), whilst some traffic does not need such a high bandwidth, resulting in low bandwidth efficiency.

2.2.2 Routing and Wavelength Assignment (RWA) in WRON

Routing and Wavelength Assignment (RWA) mechanism determines the exact link sets and wavelengths for each lightpath. There are many topics associated with the RWA mechanism in WRON. The following sections mainly focus on the wavelength requirements, and various algorithms available for RWA problem. Additionally, as the wavelength conversion is still expensive for commercial deployment, all the following description is for wavelength continuous lightpath assignment, where each lightpath has to yield a unique wavelength throughout.

2.2.2.1 Wavelength Requirements

One of the crucial topics under study is the minimum number of wavelengths required to accommodate the traffic demand.

Under the uniform traffic pattern⁸, the minimum number of wavelengths required for zero-blocking probability depends on the network topology [BB97]. For example, 13 wavelengths are sufficient to establish one lightpath for each source-destination pair in US backbone NSFnet, whilst 18 wavelengths are required for European EONnet⁹.

Under arbitrary traffic patterns, the calculation of the chromatic number involves

⁸ Uniform traffic pattern means that an equal number of lightpaths are to be established for all the source-destination pairs in the network.

⁹ NSFnet yields 14 nodes and 42 unidirectional links. EONnet yields 20 nodes and 78 unidirectional links. The topologies of NSFnet and EONnet are provided in appendix B.

Non-Polynomially (NP)-hard problems¹⁰, and is computationally expensive to determine [SM04].

2.2.2.2 Sequential RWA Algorithms

The RWA algorithm depicts the specific method of calculating lightpaths for a set of traffic demands in WDM networks. Existing RWA algorithms include sequential algorithms and synchronous algorithms.

Sequential algorithms successively decide lightpaths for each request, based on various forms of Dijkstra algorithms or pre-defined paths. For example, in k-Shortest-Path First-Fit (k-SP-FF) approach [DZB04], k disjoint shortest paths between the source and the destination are calculated offline and stored for later use. Meanwhile, each wavelength is arbitrarily assigned an index. When the lightpath is to be established, the algorithm firstly picks one of the shortest paths, and then finds a wavelength with the lowest index available in all the links of the selected path. If no wavelengths are available for the selected path, the algorithm will pick up another path until all the pre-defined paths are searched. Such an algorithm is computationally simple and fast. However, the algorithm, especially when k equals to 1, suffers from high blocking probability, because there is no flexibility in path selection.

Sequential RWA algorithms can also be flexible in both the path and wavelength assignment. For example, weight-updated Dijkstra algorithm [FBS00] uses an undirected weighted graph to represent the network topology, where the link weight usually measures bandwidth availability on the link. For each lightpath to be established the algorithm firstly runs the Dijkstra algorithm to calculate a path with the least/maximum weight, it then goes through the links along the path to find a wavelength for the lightpath. When both the path and the wavelength are successfully found, the link weights, and wavelength deployment information are updated.

The Adaptive Unconstrained Routing-Exhaustive (AUR_E) algorithm is an adaptation of the weight-updated Dijkstra algorithm [DZB04]. However, instead of using one graph, the AUR_E algorithm represents the network with W undirected graphs, where W is the number of

¹⁰ NP-hard problem refers to a class of problems, which can not be solved in a polynomial time with any deterministic approaches [GJ79].

wavelengths per link. For each lightpath to be setup, the Dijkstra algorithm is executed in each graph to produce a set of paths, among which the path with a least weight will be chosen, and the corresponding graph with chosen path is updated afterwards. Although such an algorithm is more complex than the k-SP-FF algorithm, it searches all possible routes for every request, so it achieves much lower blocking probability than the k-SP-FF algorithm.

Sequential RWA algorithms are fast and adaptable to changing traffic demands. However, they might not be able to respond to requests with an optimal or near optimal solution. In other words, sequential algorithms generate different solutions with various sequences. The best sequence is among the solution pool. A randomly selected sequence cannot guarantee that the solution is the best one [WD96].

2.2.2.3 Synchronous RWA Algorithms

Synchronous algorithms produce the whole lightpath placement concurrently. The placement is usually derived by formulating the problem in terms of ILP prescriptions with a specific objective and constraints, and the solution is produced via an ILP solver.

With various purposes, there have been multiple objective functions associated with the RWA algorithm. For instance, the Min-RWA problem assumes infinite number of wavelengths in the network, and aims to accommodate all the connection demands with the minimum number of wavelengths [BM96]. The Min-RWA is useful to confine the established lightpaths to the minimum number of wavelength¹¹. However, if the number of wavelengths is limited, and when the traffic demands are high, it is possible that not all the connections can be accommodated by the available wavelengths. In this case, Max-RWA should be considered such as to minimise the blocking probability [KS01][RS95][JMT04]. The goal of the Max-RWA is to achieve the maximum number of connections with the given number of wavelengths. Min-RWA and Max-RWA are the two mostly considered problems, though there are other objectives, such as the one targeting at the full establishment with the minimum number of hops [CKS04], and the one considering path protections [ZOM03], etc.

¹¹ Min-RWA achieves the placement with the minimum number of wavelengths. However, it is not necessarily the most economic consumption of the network resources, because the lightpaths in the solution may yield a large number of hops, occupying lots of wavelength link resources.

One of the obstacles with the ILP approach is that formulations with large number of variables and constraints may go beyond current computational capabilities. Particularly, in networks without wavelength conversion support, computational complexity dramatically increases with increasing number of wavelengths [WD96].

Generally, there are three means that can be attempted to accelerate the ILP solving time. The first approach is to reduce the ILP module size with less variables and constraints [TMP02]. The second strategy is to improve the ILP solver [HYH98]. More specifically, current ILP solver solves the integer problem by firstly relaxing the integer constraints and obtaining the optimal real solution in the first step (LP-relaxation). The solver then employs a Branch-and-Bound (BB) search algorithm to find the optimal integer solution within the real solution space. Due to the large number of integer variables and constraints, the BB search space increases exponentially with the model size. Therefore, it is critical to improve the BB algorithm such that the search space is effectively pruned, without compromising the solution optimality. Thirdly, the ILP model can be solved by relaxing the integer constraints (LP-relaxation) and then using a heuristic algorithm to deduce the integer solution [SK01]. However, such an approach may lead to a sub-optimal solution, and therefore lose the advantage that ILP approach offers.

ILP is a key approach for network optimisation, and provides benchmarks to validate and test various RWA algorithms [BM96][KS01][RS95][JMT04][CKS04][ZOM03]. However, this approach is computationally expensive. Therefore, it is very valuable to have a fast ILP-based approach to derive the optimal or near optimal solution within short time duration. More details on specific ILP formulations for Max-RWA problem will be provided in Chapter 5, where a new two-stage ILP mechanism is proposed and studied to achieve the optimal or near-optimal solution in a short time frame.

2.3 Optical Packet Switching (OPS)

OPS aims to achieve packet-by-packet forwarding, providing the maximum flexibility to the changing traffic demands. Like the packet forwarding in IP networks, optical packets are delivered based on the individual hop-by-hop header processing and switching, so the general

OPS is connectionless¹². The main advantage of OPS is that the wavelength resources are not dedicatedly reserved and they are shared by traffic from many sources. Therefore, OPS provide statistical multiplexing gain, enhancing resource utilization efficiency.

2.3.1 OPS Operations

In OPS, each optical packet contains a header and a payload, where the header and the payload are bit-rate coded and multiplexed on the same wavelength. The header and the payload can be encoded with different rate by using distinct encoding formats. For example, the header can be coded to 2.5Gbps and the payload can be coded to 40Gbps, such that the payload is highly suppressed and optically switched, whilst the header with lower bit-rate is processed by lower cost electronics [BOR00]. The packet transmission can be synchronous or asynchronous with fixed or un-fixed packet length. For example, in the European Commission supported KEOPS (Key to Optical Packet Switching) project, each packet has a fixed length and is transmitted on regulated time slots, such as to provide synchronous switching and buffering [GRG98]. Nevertheless, the packet can also be variable-length and transmitted at any time points, offering more flexibility, but requiring more control [VH05]. As there is no settled standard for OPS packet format, the packet can also be an assembled entity consisting of multiple upper-layer packets. The packet aggregation performed in edge equipments enlarges the transmission granularity, and helps to reduce the packet blocking probability [YXB02].

Once the optical packets are formed at the optical edge nodes, they are independently transmitted on the optical fibre. When entering the OPS OXCs, as shown in Fig 2-7 [YXB02], packets are firstly de-multiplexed into individual wavelengths. The packet header is then extracted from the packet and processed by a header processor. Depending on the type of the header processor used, the header can be handled optically or electronically. Meanwhile, the payload remains in the optic domain, and may have to be looped in the input Fibre Delay Lines (FDL) to wait for the completion of the header processing. After the output port is determined and configured by the header processing, the payload is optically switched to the corresponding output port via the switch fabric. At the output interface, a new packet header is generated and

¹² With MPLS/GMPLS as the upper layer control plane, an end-to-end label switched lightpath is established for each packet before the emission of the packet. This is called label switched OPS [BOR00].

attached to the packet. To alleviate network contention, wavelength conversion is required by the output interface, and the packet may be further delayed in the FDLs to resolve the contention if the output port is used by other sources¹³.

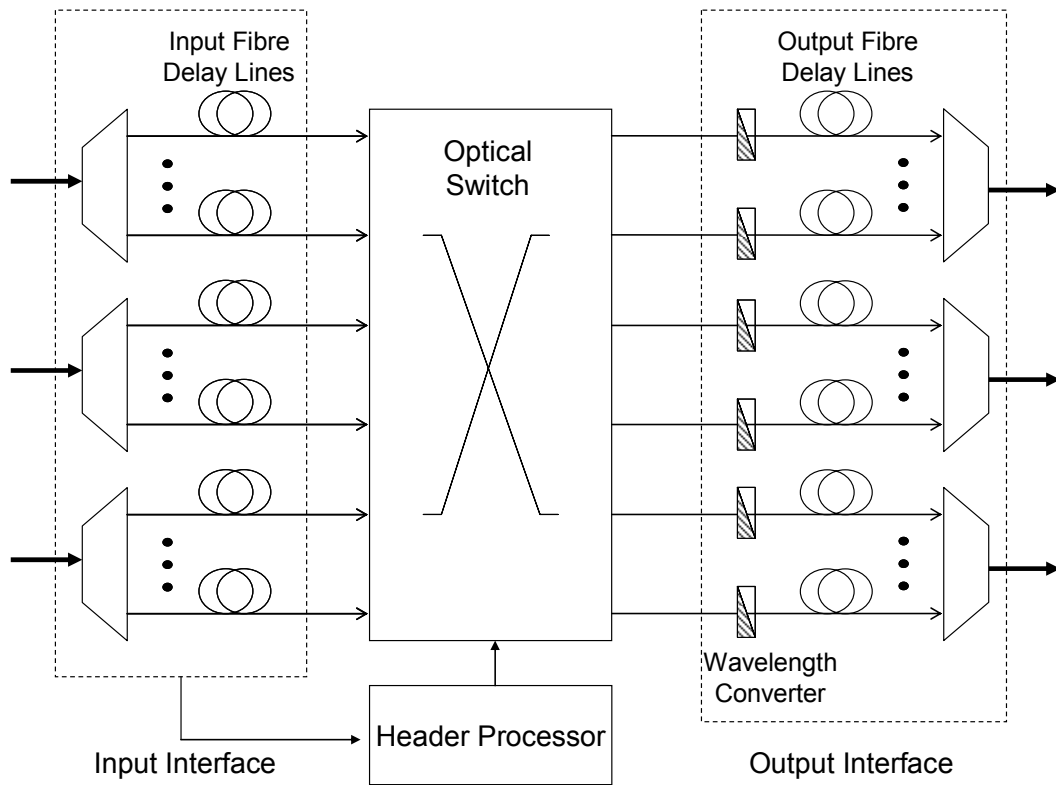


Fig 2-7: A generic structure of the OPS OXC

2.3.2 Hardware Challenges for OPS

OPS shares resource among all the traffic travelling through the network, and is adaptable to varying traffic load, but OPS is still far from the commercial deployment, due to the laggard optical device development.

Firstly, the header processing in OPS at least means to find out the output port based on the information carried in the optical packet header and the routing information contained in the routing table. Electronic header processor can perform powerful logic operations, flexible buffering, and routing table lookup and updates. However, the electronic header processor operates at about 1GHz frequencies, and is difficult to scale up to match the optical line rates

¹³ Depending on the design of the OPS switches, the FDLs buffering can also be allocated in the input interface to alleviate the contention.

[Duser03]. In optical header processor, several functions have been implemented by using special optical devices. For example, the Terahertz Optical Asymmetric Demultiplexer (TOAD) can de-multiplex a pico-second time slot from a several hundred pico-second address frame, so it supports ultra-fast header address recognition up to 250Gbps [TSP00][GSP94]. The Semiconductor Laser Amplifier in Loop Optical Mirror (SLALOM) using the two-pulse correlation principle has demonstrated the ability of optical packet header recognition at different bit-rates [DHL03]. The Tuneable Fibre Bragg Grating (FBG) technology has shown its suitability for optical routing [MK03], etc. However, the intelligence of optical header processing is very limited, not allowing for easy reconfiguration required by the packet routing. For example, it is very hard to optically hold and update the routing table information. Furthermore, the performance demonstrated by the proposed techniques is all based on the small laboratory modules. In fact, these optical devices consist of various optical elements, such as the SOA, splitters and combiners, and they are very hard to be fully integrated [Duser03].

Secondly, the time allowed for OPS switch reconfiguration is critical. Typically with the average packet length of 389.5 byte [XVC00] and 10Gbps wavelength rate, the OPS switch has to operate in nano-second regime. Additionally, the number of supported port counts per switch is also important for large scale OPS switch. In long-haul backbones, the number of fibres attached to each switch node can range from 2 to 7 [BB97], and with 128 wavelengths per fibre, the required number of port counts per switch ranges from 256 to 896. Currently, the MEMS optical switch is the most promising technology for all-optical switch, because it can be monolithically integrated on the same substrate by using the matured fabrication process of the integrated circuit industry, whilst other types of optical switches have to be cascaded from multiple 2×2 units [YLG01][YYT05]. In 2001, the Lucent Technology announced a fully integrated 3D MEMS¹⁴ optical switch with 64×64 ports and a switching time less than 10 milliseconds [YLG01]. Meanwhile, the Lucent Technology also demonstrated a 1296×1296 3D MEMS optical switch with an average switching time of 5 milliseconds in Optical Fiber Communication Conference and Exhibit, 2001 (OFC2001) [RKH01]. Generally speaking,

¹⁴ There are two types of MEMS switches called 2D and 3D MEMS, respectively. In 2D MEMS, mirrors only have two positions, ON or OFF, to redirect the lightwave or not. In 3D MEMS, mirrors can rotate in two axis, so the light can be redirected to multiple angles. 3D MEMS can reduce the number of mirrors for non-blocking switching.

current available 3D MEMS switches operate in millisecond region, and can hardly fulfil the nano-second switching speed as required by the OPS. Therefore, there is a gap between the demand and the reality in terms of the switching speed and the optical switch scale.

Thirdly, optical buffering is required in OPS to temporarily hold the payload for the completion of header processing, and to resolve network contention. FDLs are the most commonly used devices. However, the storage capacity and the storage time are limited by the fibre length, and are therefore not especially flexible. To achieve low blocking probability, a large number of lines are required, imposing integration difficulty [LT03].

Finally, OPS requires wavelength conversion to increase transmission bandwidth, and to alleviate contention. Among numerous wavelength conversion techniques reported to date, the most commercially acceptable technique is the Optoelectronic (O/E-E/O) wavelength conversion, in which the signal has to be transformed between the optic and electronic. The electronic processing directly limits the conversion speed, and a new generation of very high-speed electronics must be developed to support high-speed conversion. In addition, the cost of this type of wavelength converter is very high, for example, a 2.5Gbps optoelectronic wavelength converter costs approximately \$9000 U.S. dollars, and the cost increases with the bit-rate and the number of elements [Yoo96]. Alternatively, all-optical wavelength converters, such as the Semiconductor Optical Amplifier (SOA) cross-gain and cross-phase modulation, four-wave mixing wavelength conversion, are proposed to offer potential advantages over the optoelectronic counterpart in terms of lower costs and higher transparency. The techniques predominately rely on the nonlinear optical materials. However, the techniques suffer from various performance impairments, such as the high spontaneous noise, narrow conversion bandwidth, and the difficulty in chip integration, etc [Yoo96]. All-optical wavelength conversion is still under laboratory study, far from maturity.

As a conclusion, the technical difficulties in achieving all-optical packet networks lie in the complexity of building large, single-stage all-optical packet switches and the lack of scaleable optical buffers. Optical header processing and optical buffering will not be able to substitute electronics for the foreseeable near future.

2.4 Optical Burst Switching (OBS)

OBS is a modification of OPS [CQY04]. It enlarges the transmission unit from a packet to an aggregation of packets, called a burst, such that the header processing frequency and the required header processing speed are alleviated.

Briefly, in an OBS network, data bursts are assembled from packets at network ingress, and are disassembled back into packets at the network egress. Each burst is preceded by its own signalling message called Burst Head Packet (BHP). The BHP travels in a separate signalling channel, and slightly ahead of the data burst to reserve the path or wavelength, undergoing electro-optical conversion at every hop, whilst data burst is simply switched in all-optical manner.

More specifically, classical OBS yields three outstanding features. Firstly, the transmission unit is a burst, which may consist of one or more IP packets, a stream of ATM cells, or the raw bit streams from remote data sensors. The aggregated burst length is within the range of several kilobytes, whilst, in OPS, the transmission unit is a packet within the tens or hundreds of Bytes range. Therefore, the granularity of OBS is bigger than that of OPS, allowing more time for header processing, and reducing the burden of electronic devices that controls the optical switching configuration.

Secondly, the transmission of each burst is preceded by a BHP. BHP is usually transmitted on a separate signalling channel and converted back to electrical domain at the core node for further process, whilst the data burst is handled all optically once it enters the core. BHP is sent out ahead of the data burst by a time gap, called offset time, to inform intermediate nodes of the upcoming data burst, so that they can make a routing decision and configure their fabric to switch the burst to the appropriate output port. Sufficient offset time puts buffering at the network edge, making it possible to transmit data burst in all-optical manner without requiring a big optical buffering at the intermediate node.

Thirdly, in most of the existing schemes, OBS is based on the one-way reservation, and a data burst is emitted into the core without the knowledge of end-to-end connectivity [CQY04][WM00]. Although this can help avoid a round-trip-time delay, the burst may be

dropped at the intermediate node in case of congestion or output port conflicts. Thus, OBS is best-effort-service as in connectionless packet switching, without guarantee of delivery. High blocking probability is the major disadvantage of the OBS technology.

In the following sub-sections, various important aspects of OBS networks, such as the structure of the edge/core OBS nodes, the burst assembly mechanism, and the signalling protocols, etc. are introduced and explained in detail.

2.4.1 OBS Operations

Fig 2-8 provides a typical structure of OBS networks, where the inset demonstrates the separated transmission of bursts and the corresponding BHPs.

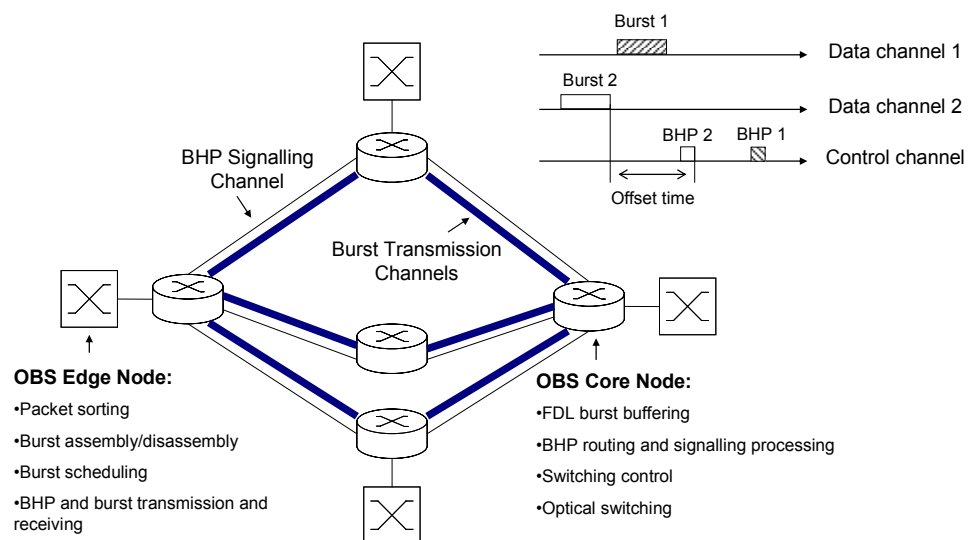


Fig 2-8: OBS network structure

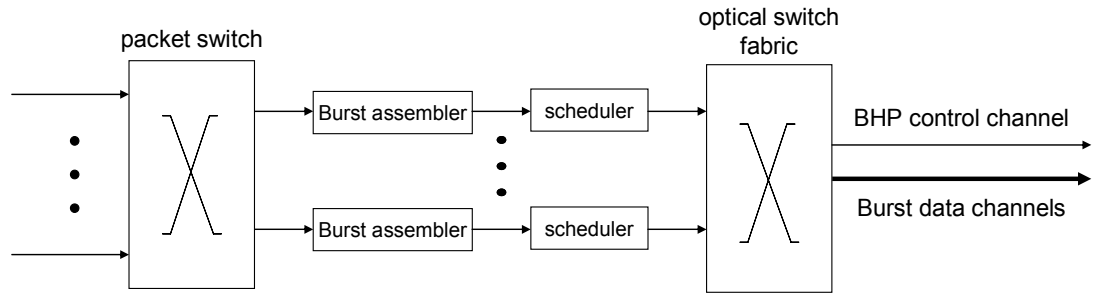
The network mainly consists of OBS edge nodes, and OBS core nodes. The OBS edge node is the entity containing most intelligence and complexity, and is mainly in charge of burst assembling/disassembly, burst scheduling, and BHP manipulation, etc. The OBS core node deals with the BHP processing, and sets up the switch fabric for the upcoming burst.

2.4.1.1 OBS Edge Node

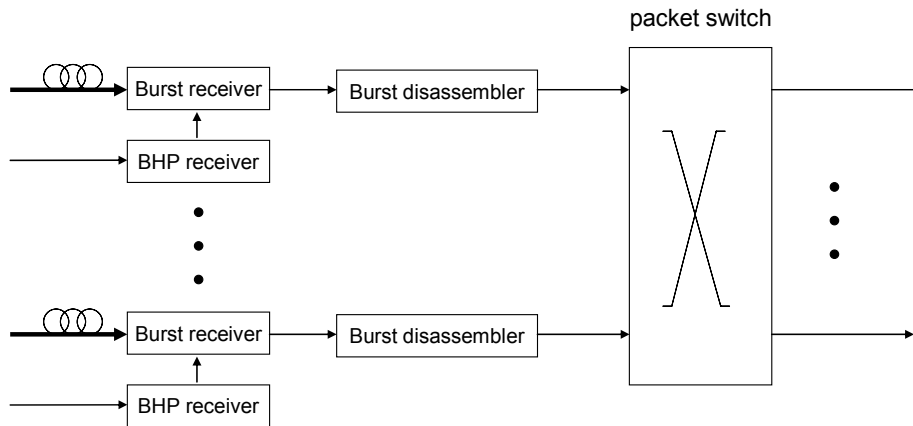
OBS edge node is situated at the edge of the core, collecting packets from legacy interfaces, such as IP/ATM, Ethernet, etc, and manipulating edge functions mainly in electronic domain.

Fig 2-9 presents a general structure of the OBS edge node for both the sending and the receiving

sections [XVC00][ROB04].



A: Sending section of the OBS edge node



B: Receiving section of the OBS edge node

Fig 2-9: General structure of the OBS edge node

In the sending section, packets from the upper layer are firstly sorted according to the destination and the class of service, and the packets are then statistically multiplexed and switched to the corresponding burst assemblers via the packet switch. In the burst assemblers, packets are collected and assembled into a burst according to a specific burst assembly mechanism, which will be described in a later section. The scheduler keeps tracks of unscheduled time slots in both the data channels and the signalling channels, and decides the time that the burst and its BHP can be transmitted. Once the BHP is scheduled to the core, an offset time is maintained between the BHP and the corresponding burst, after which the burst is emitted to the core. The scheduled burst and BHP are transmitted via the optical switch fabric.

In the receiving section, the functionalities are the reverse of the counterparts in the sending section. Bursts and BHPs are received by the corresponding receivers, whilst, some FDLs can be provided before the burst receivers to temperately buffer the bursts, providing time for BHP

processing and offering burst disassembly instruction. In burst disassemblers, bursts are separated into packets. Additionally, the burst disassemblers can also deal with packet re-ordering and retransmission, if required. The disassembled packets are then forwarded to their next hop via the traditional packet switch.

2.4.1.2 OBS Core Node

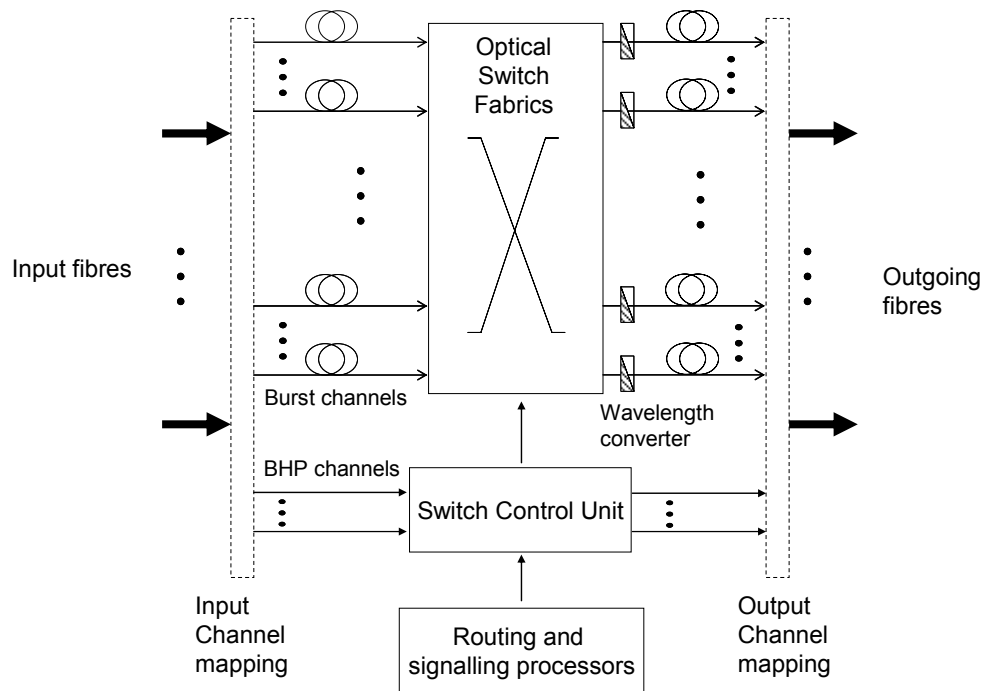


Fig 2-10: General structure of the OBS core node

OBS core node deals with the BHP routing and signalling in the electronic domain, and the burst forwarding in optical domain. The general structure can be illustrated in Fig 2-10, where the OBS core node mainly consists of the optical switch fabrics, Switch Control Unit (SCU), and the routing and signalling processors [XVC00].

When the bursts and the BHPs enter the OBS core node, the input channel mapping block decouples the traffic into burst wavelength channels and the BHP wavelength channels, where the burst channels are connected to the optical switch fabrics, and the BHP channels are connected to the SCU.

In some core node design, FDLs are provided for the input burst channels, such that the bursts can be delayed to wait for the completion of BHP processing. However, FDLs can also be completely removed from the core node by assigning a sufficient long offset time [CQY04]

[WM00]. Once the switch is ready, the burst is switched and forwarded via the optical switch fabrics in optical domain. At the output ports, wavelength converters are provided to enable wavelength conversion and expand wavelength availability. Additionally, output FDLs can help to relieve the contention when multiple bursts are simultaneously competing for the same output ports.

In the BHP control part, the routing and signalling processors maintain the routing table and provide routing decision for the SCU. The SCU then decides the specific channel that the associated burst should be forwarded to, and provides the instruction to reconfigure the optical switch fabrics. The output channel mapping block combines the bursts and the BHPs into the outgoing fibres.

2.4.2 Burst Assembly

Burst assembly mechanisms are used to regulate the burst generation. Intuitively, burst assembly enlarges the transmission granularity, such that fewer header processing takes place in the core nodes.

Basic burst assembly mechanisms include the timer-based approach, the threshold-based approach, and their mixtures [UE04][YLC04][TAK03].

In the timer-based approach, a timer is triggered when the first packet of the burst arrives at the burst assembly queue. The burst is regarded to be ready when the timer runs out. In the threshold-based approach, a burst aggregation completes when the burst length reaches the threshold.

There are advantages and disadvantages associated with both assembly schemes. In the timer-based scheme, the assembly delay is bounded by the timer counts, being easy to control the edge delay. However, the burst length can be excessively long under a high load, causing congestion and loss in the core. Similarly, the burst length can also be too short under a light load, putting stress on BHP processing. In the threshold-based scheme, the burst length is stationary. However, it might take a very long time to reach the specified burst length under a low incoming rate, increasing edge delay, or, it might release the bursts too quickly under a high load, exceeding the core switching capability.

To improve the performance of the timer-based and the threshold-based assembly schemes, a mixture burst assembly scheme is proposed, where a new control parameter, the minimum burst length B_{min} , are introduced. The two traditional schemes with the burst length threshold B , and the timer count T parameters, are combined to control the burst aggregation [YLC04]. Basically, the burst assembly finishes when either the B or the T threshold is reached. Therefore, when the traffic load is very low, and the time takes to reach the B is much longer than the T parameter, a burst assembly stops when T time is elapsed. When the traffic load is very high, and the burst length after time T is much longer than the B parameter, the burst aggregation will stop when the length touches the B parameter. Additionally, when the traffic load is extremely low, and the burst length is shorter than the B_{min} after time T , the burst will be padded to B_{min} , such as to guarantee the transmission granularity. Such a mechanism yields the lowest assembly delay, when it is compared with other schemes [ZB03].

2.4.3 OBS Signalling Protocols

Signalling protocols specify the procedure by which the network resource is reserved and utilised. The OBS signalling protocols can be classified into two categories. The first category represents end-to-end connectivity un-acknowledged protocols [CQY04][WM00], and the second category represents end-to-end connectivity acknowledged protocols [DB02][DB022][SST03][STS04].

In the end-to-end unacknowledged protocols, a burst is emitted to the core without any knowledge of end-to-end connectivity. The burst's set off time is regulated by the *offset* time parameter after its BHP starts off. However, the burst does not wait for any form of lightpath acknowledgement, being ignorant of the end-to-end connectivity. As the burst does not wait for a signalling round-trip-time, this type of protocol offers low latencies. However, in the optical core with very limited and inflexible optical buffering, some bursts will be abandoned at congested core nodes. Therefore, this type of reservation protocol suffers from an un-predictable blocking probability in the optical core nodes, which triggers the retransmission of TCP sessions, consequently causing long delays and high level of network congestion [SLK05]. Typical examples of end-to-end connectivity unacknowledged protocols include the Just-Enough-Time (JET) [CQY04] and the Just-In-Time (JIT) [WM00] protocols, which mainly differ in how long

should be the offset time between the BHP and the corresponding data burst, and also how soon before the burst arrival and how long after its departure the switching elements are made available to other bursts.

In the end-to-end connectivity acknowledged reservation protocols, each burst is preceded by a two-way acknowledged reservation signalling, so an end-to-end lightpath is established for each burst before the burst launches to the core. However, as the burst has to wait for a signalling round-trip-time, the burst will be delayed for long. Typical instances of this type of protocols include the Wavelength-Routed OBS (WR-OBS) [DB02][DB022] and Multi-Protocol Label Switched (MPLS)/Generalized MPLS (GMPLS)-based OBS [SST03][STS04] ¹⁵.

2.4.3.1 JET Protocol

JET is a one-way reservation protocol, in which no acknowledgement signalling takes place. The BHP is set off ahead of the burst to reserve the resource along the path towards the destination. Meanwhile, the burst remains at the source, and is delayed for offset amount of time before it is launched.

As indicated by its name, the offset time in JET is set to a value such that there is just enough time for the BHP to be transmitted and processed before the burst data arrives at the destination [CQY04]. Let T_{proc} denote the processing delay of the BHP at each node, and T_{sw} denote the time to configure the optical switch matrix. The value of the offset time can be formulated as in Equation 2-1.

$$T_{offset}^{JET} = \sum_{path-nodes} T_{proc} + T_{sw} \quad \text{Equation 2-1}$$

The JET offset time can also be illustrated in Fig 2-11, where two intermediate nodes are involved between the source and the destination. As we can see, the offset time is set to be the value of having just enough headroom for the destination node to complete the processing of the BHP before the first bit of the burst arrives. So it is a minimised time, which will reduce the end-to-end delay. If the offset time is less than that, the burst may bypass the BHP and be dropped.

¹⁵ As WR-OBS is usually regarded as a new type of switching technology, we discuss WR-OBS separately later.

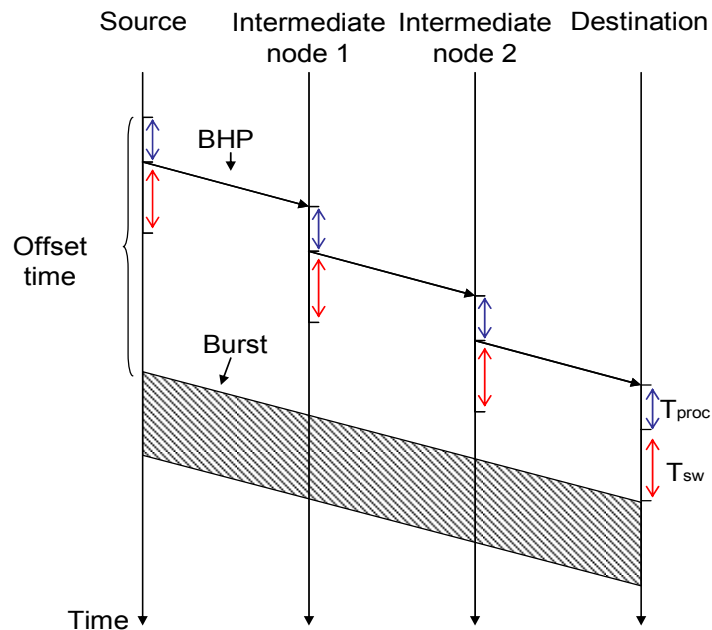


Fig 2-11: Offset time in JET protocol

In terms of the resource reservation, JET uses Delayed Reservation (DR), and assumes that BHP knows the length of the burst. The resource reservation does not start immediately after the processing of control packet at each node, but start from the time when the burst is estimated to arrive. The resource will be reserved for a duration that allows the whole burst to be transmitted. This requires BHP to carry the information of burst length so that it can specify when and how long the resource will be reserved. DR is better than immediate reservation, because the resources are reserved only when they are needed. However, it will add more complexity to the software components, and its performance greatly depends on whether the information carried by the BHP is matching the real burst.

2.4.3.2 JIT Protocol

JIT is a two-way protocol with upstream SETUP messages and downstream CONNECT messages confirming the end-to-end connectivity. However, the burst is sent out to the core without waiting for the arrival of the connection acknowledgement [WM00].

Fig 2-12 shows details of the JIT signalling procedure. When the burst is ready, a SETUP message is sent from the origination towards the destination to request a connection. The BHP processing time is denoted as T_{proc} , during which the routing table will be consulted at each hop along the path to determine the next hop and wavelength to use. The cross-connect will then be

configured using time T_{sw} .

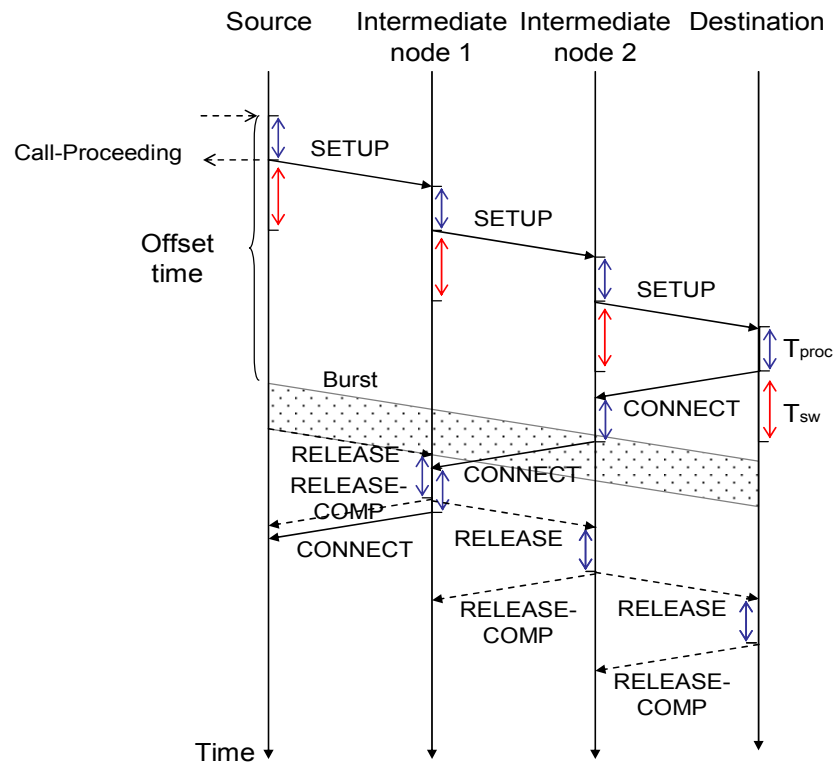


Fig 2-12: JIT signalling protocol

At the ingress node, once the BHP is processed, a Call-Proceeding message is sent back to the burst, informing the estimated offset time that the burst should wait before it is emitted to the core. The offset is calculated based on the hop number towards the destination, and the associated cross-connect configuration time.

When the SETUP message reaches the destination, a CONNECT message is sent in the reverse direction, indicating the end-to-end reach-ability. However, the burst does not wait for the CONNECT message. Instead, it is set off when the offset time duration provided by the Call-Proceeding message elapses. Therefore, the burst is still prone to being blocked during the transmission if congestion happens in some intermediate nodes.

Finally, when the transmission completes, the RELEASE message is sent in either the upstream or the downstream direction to release the resources. The RELEASE-COMP message is sent back to the RELEASE message sender, confirming that the resource has been successfully released.

In more detail, the duration of the offset time in JIT protocol is determined by the requirement that the cross-connect must be readily setup before the arrival of the burst. In addition, the destination should also complete the configuration before the burst arrives. Therefore, the JIT offset time can be expressed as in Equation 2-2.

$$T_{offset}^{JIT} = \sum_{path-nodes} T_{proc} + T_s \quad \text{Equation 2-2}$$

Where $T_s \geq T_{sw}$, is the delay adjustment to insure that the cross-connect is configured at the last node.

The resource reservation period in JIT is longer than the JET protocol. The resource is set to be held once the BHP is processed and the cross-connects are configured. However, JET uses Delayed-Reservation, where the resources start to be held when the burst arrives. Ref [DGS01] evaluates and compares the performance of JET and JIT. The result shows that JET yields the lowest blocking probability, as expected. However, JET achieves the lower blocking probability at a price of higher complexity at the scheduler in each OBS node, whilst, JIT avoids complicated software scheduling.

2.4.4 Routing and Wavelength Assignment (RWA) in OBS

OBS requires a proper access to the routing and wavelength assignment functions. Normally, the routing and wavelength assignment are processed separately.

In the routing part, the routes are derived from associated routing protocol, such as the Open Shortest Path First (OSPF) protocol. The information concerning the network and link utilization is updated and stored in a link-state database at each node. Each node will then use the link-state database to calculate the best path to other nodes in the network. The calculated routing information is stored in the routing table. In MPLS/GMPLS, the corresponding routing protocol can be enhanced to support traffic engineering. For example, in the link-state database, apart from the link length information, other traffic engineering metrics like the reserved bandwidth, maximum bandwidth can also be stored. These extra metrics can be used by the path calculation such that not only the shortest path but also the path with the least congestion can be found.

Wavelength assignment determines which wavelength to use along the path. There can be

wavelength converter enabled and wavelength converter disabled schemes. In the former scheme, wavelength converter is required at each OBS node, and the burst can use different wavelengths along the path. In the latter scheme, the burst has to use a unique wavelength along the entire path. The wavelength converter enabled scheme increases the bandwidth, so it is able to reduce the probability of blocking. However, the current state of the technology is still not applicable for full-range wavelength conversion. The wavelength converter disabled scheme yields less flexibility in wavelength selection, so it causes high blocking probability. In order to achieve low blocking probability, wavelength conversion is assumed in most of the existing OBS schemes.

In addition, in JET or JIT, bandwidth on each wavelength channel is fragmented by bursts with some void intervals. If the void bandwidth is not utilized, lots of resource is wasted. Therefore, a good wavelength scheduling algorithm associated with the wavelength assignment is required to improve the bandwidth utilization and reduce the blocking probability. Latest Available Unused Channel with Void Filling (LAUC-VF) is a well-known scheduling algorithm that achieves high bandwidth utilization [XVC00][XQL04]. LAUC-VF keeps track of all void intervals on the wavelength channels. When a burst is expected to arrive at time t , the LAUC-VF will search the wavelength channels to find the one with a large enough interval, and whose latest available time is the closest to the arrival time of a new burst.

2.4.5 Congestion Resolutions in OBS

The end-to-end connectivity un-acknowledge OBS protocols, such as the JET and JIT, suffer from resource contention, where bursts have to be abandoned if two or more bursts are destined to go out of the same output port at the same time. To date, optical buffering, wavelength conversion, deflection routing, and burst segmentation are four main technologies to resolve contention.

Currently, FDL is the only optical buffering available with today's technology. The optical data is buffered in FDL by simply circulating in the FDL. In other words, the optical data is temporally delayed; the delay depends on the length of the line. Yoo et al. [YQD00] studies the effect of the FDL, and proves that FDL can reduce the blocking probability. However, FDL is not scalable enough, and is not commercially viable.

Wavelength conversion can relieve contention by sending bursts through different wavelengths. However, it is still an expensive technology.

In deflection routing [KKK02], when there is a conflict between two packets, one packet is switched to the destined port, while the other is switched to any other available port. In this way, no or little buffering is needed. However, the end-to-end delay for a packet may be unacceptable, and the packets may be out of order.

In burst segmentation, rather than dropping the entire burst during contention, the burst is broken into several segments, and only overlapped segments are dropped. Ref [RH03] evaluates the advantage of the segmentation on the blocking probability both in mathematics and simulation. Ref [VJS02] also investigates the segmentation scheme in conjunction with the deflection routing, and the result shows that segmentation with deflection can achieve significantly reduced packet loss rate.

2.4.6 QoS Provisioning in OBS

End-to-end connectivity un-acknowledged OBS only provides best effort service, but mission-critical and real-time applications require a QoS with specified low delay and loss probability.

Typically, Chunming Qiao et. al. propose a priority JET (pJET) [YQ98][YQ99][YQD00], where bursts are classified into multiple classes, and differentiated services are provided. The basic idea of pJET is to assign an extra offset time to the high priority burst. Therefore, this is also known as offset-time based scheme. The extra offset time allows a BHP to make resource reservation in much more advance, thus giving a greater chance of success. However, the burst with high priority will have longer offset time and delay than a low priority burst.

There are several other QoS mechanisms. However, the common problem is that these mechanisms will have the effects of reducing the burst loss for high class traffic, whilst at the expense of increased burst loss for low class traffic.

2.5 Wavelength-Routed OBS

WR-OBS is a type of OBS that requires mandatory end-to-end acknowledgements. WROBS

combines OBS with a fast circuit-switching dynamic lightpath assignment mechanism to provide end-to-end guaranteed services without requiring wavelength conversion support [DB02][DB022][DZB04][MKB02]. As the WROBS employs the two-way reservation scheme to dynamically setup and release lightpaths as required, it enlarges the transport unit from microsecond-small bursts to millisecond-large bursts. Though the two-way reservation requires round-trip-time delay, it guarantees a deterministic end-to-end delay for each burst, and the burst will not suffer from any loss during the transmission and propagation.

2.5.1 Architecture of WR-OBS Networks

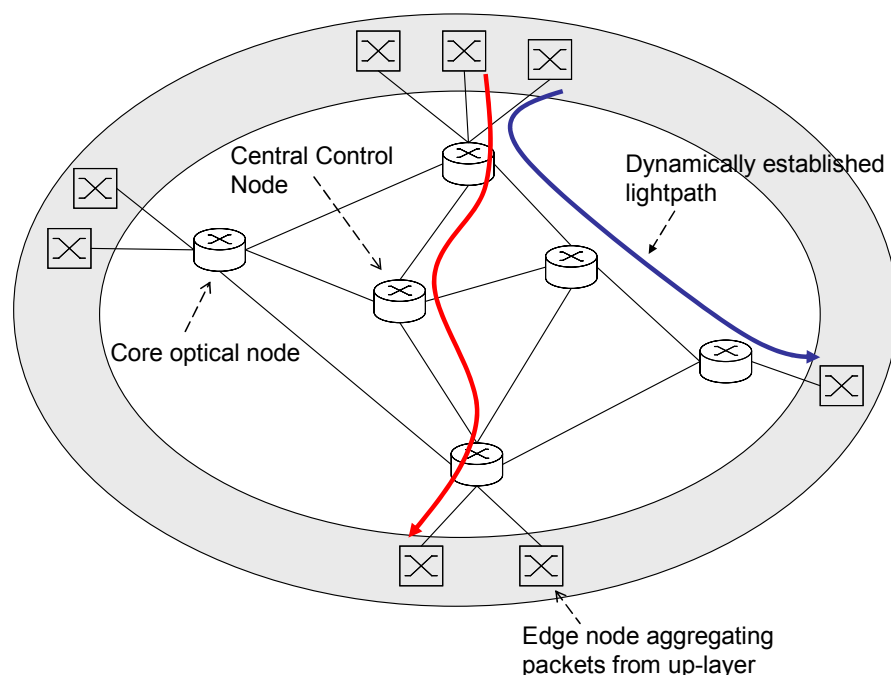


Fig 2-13: WR-OBS network architecture

Fig 2-13 shows the architecture of WR-OBS networks. Briefly, the current WR-OBS study adopts centralized control architecture. Packets are electronically aggregated into bursts at the network edge according to their destination and class of service. The aggregation time is within millisecond region, and it is determined by the performance parameters such as the edge delay or the required burst size for the network. At an appropriate point during the aggregation cycle, an end-to-end lightpath is requested from the network control node. Once a free lightpath is found, the corresponding OXC is configured, and the reservation acknowledgement is sent back to the ingress node. The burst is assigned to the lightpath and is transmitted into the core

network. The burst's further latency only depends on the propagation delay. Following the transmission, the lightpath is released and can be reused for subsequent connections. WR-OBS puts all the processes and buffering at the network edge, enables a bufferless core network, and simplifies the OXC structure.

2.5.2 Burst Aggregation Mechanisms in WR-OBS

There are two burst aggregation mechanisms proposed for WR-OBS networks, called Limited-Burst-Size (LBS) and Unlimited-Burst-Size (UBS) scheme, respectively. Both of the aggregation schemes are operated with a sense of “reserve-while-accumulating” behaviour.

Particularly, in LBS scheme, the burst aggregation starts when the first bit of the new burst arrives at the assembly queue. After a time-out signal indicates the potential buffer overflow or the potential violation of end-to-end tolerance, the lightpath reservation request is sent to the control node to ask for the reservation of a lightpath. Meanwhile, the burst assembly goes on until a lightpath acknowledgement is received by the edge node. Any new packets arrive later than the receipt of the acknowledgement will be regarded as a part of the next new burst.

The UBS scheme is similar to LBS scheme in many senses. However, when the lightpath acknowledgement is received by the edge node, the burst assembly does not cease. Instead, the existing packets in the assembled burst start to be transmitted. Meanwhile, packets newly arriving during the transmission is counted as a part of the current burst, and the aggregation keeps going until the assembly queue becomes empty. Existing simulation studies show that UBS yields better performance than the LBS in terms of blocking probability [DB02][MDB01].

2.5.3 QoS Provisioning in WR-OBS

In most of the classical OBS schemes, the low burst loss for high priority traffic is achieved at the expense of losing more low priority traffic bursts. In WR-OBS, a different QoS provisioning method is used, such that no compromise happens to the burst blocking probability for all classes of traffic, whilst keeping end-to-end delays within the bounds as specified by each class of traffic [KDZ02][KB02].

The QoS differentiation in WR-OBS is achieved by the request server in the control node, as shown in Fig 2-14

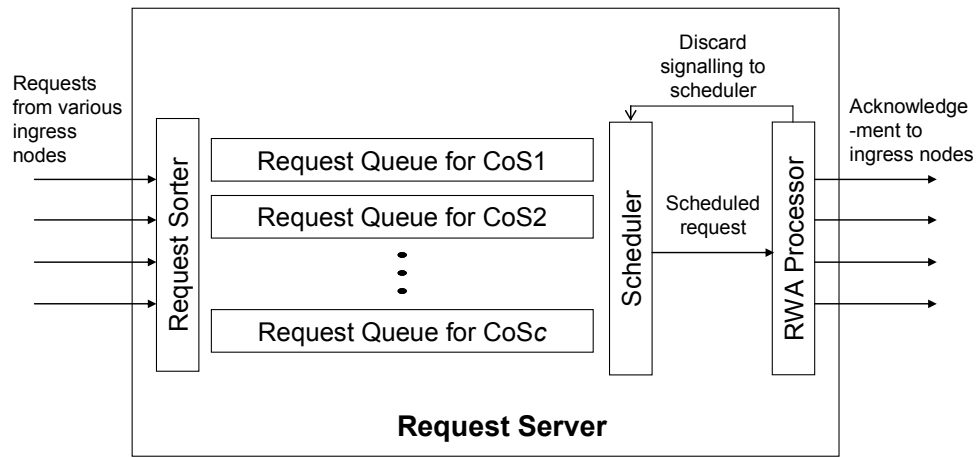


Fig 2-14: Request server in WR-OBS control node

More specifically, at the edge node, there are $C(N-1)$ burst assembly queues, representing C classes of services for $(N-1)$ destinations. Packets are stored in the corresponding assembly queue according to their CoS and destination.

In the control node, there are C request queues representing C CoS levels. The requests are sorted according to their CoS, and are directed to their dedicated queue. The mechanism for service differentiation consists of two functions. The first function is a fair queuing algorithm for the queue sequencing. For example, a weighted round-robin algorithm (WRR) can be used such that the sequence of queue processing is determined by an assigned weight for each queue. The second function is a scheduling principle for the request sequencing inside each queue. For example, the EDF principle can be employed such that request with the earliest deadline is scheduled first.

The QoS provisioning is also associated with the end-to-end delay. The end-to-end delay is the time elapsed from the start of the burst aggregation until the first bit of the burst reaches the destination. In WR-OBS, the end-to-end delay is controlled by a parameter, called $T_{sched,max}$. $T_{sched,max}$ is the maximum time allowed for the requests of the class C_i to be buffered at the request queue in the control node. If no lightpath is found for the request when $T_{sched,max}$ is exceeded, the request is discarded and its burst is blocked.

Simulation results for WR-OBS with multiple CoS levels show that the request scheduling allows for all traffic classes to keep the same burst blocking probability, at the expense of the increased end-to-end delay for low class traffic. This can be easily achieved by assigning a

longer $T_{sched,max}$ for low class traffic. Furthermore, to achieve the specified burst blocking probability, the value of required $T_{sched,max}$ is affected by the number of wavelengths in the network. The more wavelength does the network support, the shorter the $T_{sched,max}$ has to be, hence the shorter the end-to-end delay.

2.5.4 Provision of End-to-End Delay Guarantees in WR-OBS Networks

Currently, more and more new optical services, such as online multimedia services, emerge in the network. These services are not only loss sensitive but also latency sensitive. In WR-OBS, bursts experience a round-trip-time delay by the lightpath signalling before it is transmitted. Therefore, the newly emerged services are relatively sensitive and constrained by the lightpath setup delay.

As mentioned in the previous subsection, the end-to-end delay in WR-OBS networks is controlled by $T_{sched,max}$, the total end-to-end delay for each burst can be expressed as in Equation 2-3 [MKB02].

$$\Delta_{ee} = T_{aggr} + 2T_{sig} + T_{rwa} + T_{sched,max} + T_{tuning} + T_{prop} \quad \text{Equation 2-3}$$

Where T_{aggr} is the time-out value for initial burst assembly, T_{sig} is the signal propagation time from the edge node the control node, or versa vice. T_{rwa} is the time required for executing the RWA algorithm. T_{tuning} is the time required for tuning the laser in the source node and for the reconfiguration of the OXC along the lightpath.

Simulation results with strict requirements on end-to-end delay, bounded by 90ms, and on blocking probability 10^{-4} , demonstrate that a traffic load¹⁶ of 0.35 can be supported using 12 wavelengths in NSFnet. However, if a lower number of wavelengths, such as 10 wavelengths, is used, the maximum traffic that can be supported is dramatically reduced to as low as 0.01.

The reason for this is related to the low lightpath bandwidth efficiency in setting up the lightpath in large backbones. Specifically, the lightpath resources are set to be held during the lightpath

¹⁶ Traffic load is defined as the ratio of incoming traffic bit rate to the outgoing wavelength rate. In ref [MKB02], the same amount of traffic load is generated for each source-destination pair.

acknowledgement phase. In large backbones, the acknowledgement propagation time is long, so lots of resources are unnecessarily held in an idle state during this period time, reducing bandwidth efficiency. Therefore, the simulation with a 1/3 original NSFnet scale illustrates that higher traffic loads can be supported in such a reduced network.

2.6 Summary

Over the last decade, SONET/SDH systems have become the *de facto* standard for backbone optical networks. WDM/DWDM has been introduced to boost the capacity in a single fibre. However, WDM/DWDM is only used as a transmission media within current SONET/SDH structure, and such a transmission system does not operate optimally. Though progress has been made towards AON, there are still no possibilities of commercial deployment in near future.

Future optical network architecture design aims to achieve more flexible and economical use of network resources with eliminated OEO conversion. Meanwhile, it is also important to deliver satisfactory services in terms of the blocking probability and the end-to-end delay. WRON, OPS, OBS, and WR-OBS are potential network architectures for next generation optical networks. However, none of these propositions are perfect, so further efforts are required to improve the network performances, such as the blocking probability performance and the ability of supporting latency and loss sensitive traffic.

Chapter 3 Existing Work on Predictive Reservation Mechanisms

This chapter describes existing work on predictive reservation mechanism, mainly the Forward Resource Reservation (FRR) scheme, which makes use of traffic prediction to reduce the latency and provide QoS differentiation in the classical OBS networks [LN02][LN03][LNJ03]. This chapter also introduces time-series models for traffic prediction, and finally outlines the research focus of this thesis.

3.1 Forward Resource Reservation (FRR) Mechanism

So far, a FRR mechanism that makes use of traffic prediction technology has been proposed to achieve the latency reduction and QoS provisioning in classical OBS networks [LN02][LN03][LNJ03].

In classical OBS, the transmission of the BHP is initialled after the burst assembly process. In other words, to acquire the necessary information of the corresponding burst, a BHP waits for the completion of burst assembly before transmission is started. Each BHP reserves network resources based on information such as the burst length and the expected burst arrival time. The burst is then sent off to the core after the expiry of the offset time.

In FRR, latency reduction is achieved by processing the burst assembly and the BHP transmission in parallel, and thereby minimizing their impact on the end-to-end delay.

As shown in Fig 3-1 [LNJ03], as soon as a new burst assembly begins at time T_b , the ingress edge node predicts the length of the burst in progress, following which the BHP is constructed with predicted burst length information. Instead of waiting for the completion of the burst assembly, the BHP is sent to the core at time $T_h = \max(T_b, T_b + \tau_a - \tau_o)$, where τ_a is the burst assembly time, and τ_o is the offset time.

When the burst assembly completes at time $T_b + \tau_a$, the actual burst length is compared with the reservation length in the pre-transmitted BHP to ensure the pre-reserved duration is enough for the actual burst length. If the actual burst length is no longer than the pre-reserved duration, the

BHP pre-transmission is deemed as a success, and the burst will be sent at time $T_h + \tau_o$. However, if the actual burst is longer than the pre-reserved duration, the BHP pre-transmission fails to provide sufficient resources for the burst, and is deemed invalid. The BHP has to be retransmitted for this burst at time $T_b + \tau_a$ with the actual burst size, and the burst lags behind by the offset τ_o .

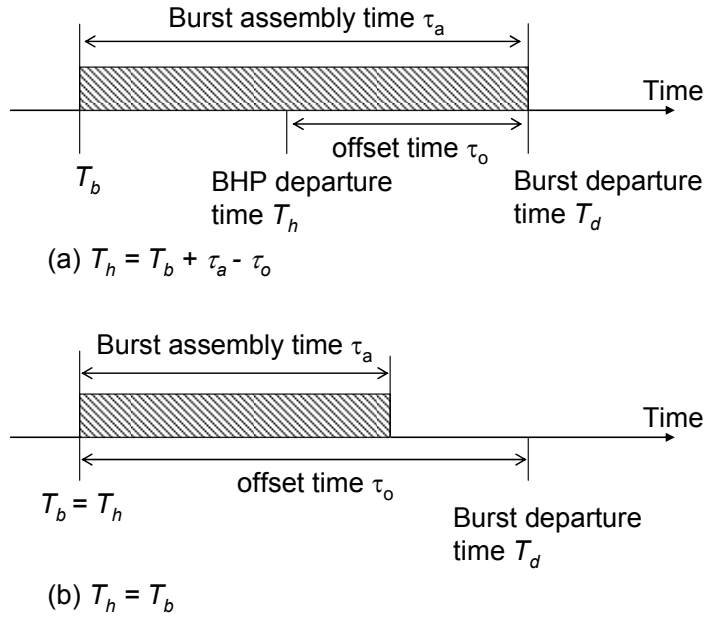


Fig 3-1: FRR mechanism

3.1.1 Aggressive Resource Reservation

The FRR mechanism adopts an aggressive resource reservation strategy to determine the resource reservation duration.

More specifically, due to the imperfections in the predictor, the estimated length may be longer or shorter than the actual burst length. In the case that the predicted burst length is shorter than the actual burst length, the pre-reservation fails and the BHP has to be sent again to make the reservation for the actual burst length. This not only wastes the efforts of the pre-reservation process, failing to achieve any latency reduction, but also wastes pre-reserved resources¹⁷.

¹⁷ The resource wastage caused by the insufficient pre-reservation can be avoided by Bandwidth Enhanced FRR mechanism, as in ref [LN03]. Basically, the edge node monitors the actual burst length during the burst assembly. As soon as the actual burst length exceeds the predicted length, a CLEANUP message is sent along the way to cancel the pre-reservation, and make the resource available for other bursts.

In order to improve the pre-reservation success probability, the FRR mechanism adopts an aggressive resource reservation method, where a small margin of correction δ is introduced to the predicted burst length. In the end, the pre-reserved burst duration is the predicted burst duration L_p plus the correction δ as expressed in Equation 3-1.

$$L_{pre-reserved} = L_p + \delta \quad \text{Equation 3-1}$$

If the prediction error is zero-mean Gaussian noise with variance equal to σ^2 , the correction value δ is controllable by setting it as multiples of σ . Thus, the probability that the pre-reservation successes can be easily expressed by the Q-function¹⁸. For example, by setting the correction value δ three times of the σ , the probability that the pre-reservation fails is only 0.135% [LNJ03].

Increasing the magnitude of the resource reservation enhances the pre-reservation success probability, and makes the FRR more efficient in terms of latency reduction. It is shown that if the burst length can be precisely predicted, such that the pre-reservation succeeds with a very high probability near 100%, the FRR scheme can reduce the edge node latency by 66%.

3.1.2 FRR for QoS Provisioning

To achieve a flexible QoS differentiation for different classes of applications, the FRR scheme can be adjusted to support QoS provisioning [LN02][LN03]. The aim is to reduce the burst delay for the latency sensitive applications, such as interactive video conferencing.

The service differentiation is mainly carried out by assigning each class of traffic different value of τ_p to control the time when to launch the BHP prior to the burst assembly completion [LN02], as shown in Fig 3-2. Assuming that there are two classes of service, the class 0 service is latency sensitive, and the class 1 service is latency insensitive. As shown in Fig 3-2, for class 1 traffic, a simple resource reservation is executed, where the BHP is generated and sent to the core after the completion of the burst aggregation. For class 0 traffic, the FRR mechanism is adopted, and the BHP is emitted at a time of τ_p prior to the completion of burst assembly. The delay of the

¹⁸ Q-function is defined as the probability that a standard normal random variable (zero mean, unit variance) exceeds x. The Q-function is a convenient way to express right-tail probabilities for normal (Gaussian) random variables.

class 0 traffic is thus reduced. The parameter τ_p is adjustable, so for more classes of service, different value of τ_p can be specified by the user/operator as a QoS constraint.

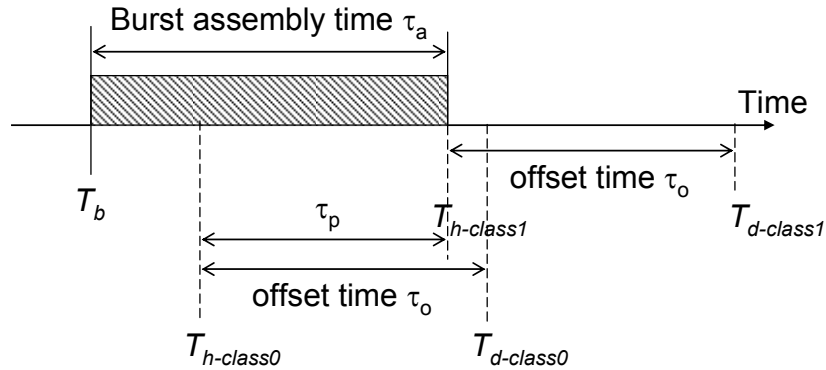


Fig 3-2: QoS provisioning in FRR mechanism

3.1.3 Linear Predictive Filter (LPF) for Traffic Prediction

In FRR, traffic prediction is carried out by a LPF, where the LPF is based on the Auto-Regression (AR) Model in time series theory [LN02][LN03][LNJ03][Adas98].

In the AR model, time series is a stationary process with constant mean value μ and variance ν . The current value of the process is expressed as a linear combination of previous p values, plus a random noise component, where p is called the order of the AR model.

More specifically, let's denote the values of a process at equally spaced times $t, t-1, t-2, \dots$ by $Z_t, Z_{t-1}, Z_{t-2}, \dots$, let $Z_t', Z_{t-1}', Z_{t-2}', \dots$ be deviations from mean value μ , that is, $Z_t' = (Z_t - \mu)$, the Z_t' can be expressed as in Equation 3-2, where a_t is a random Gaussian variable with zero mean and variance ν' , and ϕ_i is the AR coefficients. Correspondingly, the value of Z_t can be expressed as $Z_t = Z_t' + \mu$.

$$Z_t' = \phi_1 Z_{t-1}' + \phi_2 Z_{t-2}' + \dots + \phi_p Z_{t-p}' + a_t \quad \text{Equation 3-2}$$

Applying the AR model in LPF, the process is assumed to yield a zero mean, the predicted value at time $t+1$, can be expressed in Equation 3-3.

$$Z_{t+1}^{predicted} = \sum_{i=0}^{p-1} (\phi_i \times Z_{t-i}) \quad \text{Equation 3-3}$$

In LPF, the AR coefficients can be calculated by at least two approaches. The first approach is

the Yule-Walker method, and the second approach is the recursive Least Mean Square (LMS) method. Yule-Walker equations rely on the auto-correlation function of the observed process to derive the AR coefficient.

3.1.3.1 Yule-Walker Equations for AR coefficients Estimation

Yule-Walker equations rely on the auto-correlation function of the observed process to derive the AR coefficient. The deduction of the Yule-Walker equation is rather simple. Let's multiply Z_{t-k}' to both sides of Equation 3-2, we get Equation 3-4.

$$Z_{t-k}'Z_t' = \phi_1 Z_{t-k}'Z_{t-1}' + \phi_2 Z_{t-k}'Z_{t-2}' + \dots + \phi_p Z_{t-k}'Z_{t-p}' + Z_{t-k}'a_t \quad \text{Equation 3-4}$$

By taking expected values in Equation 3-4, we can obtain Equation 3-5, where γ denotes the auto-covariance function, with $k>0$, and $E[Z_{t-k}'a_t] = 0$.

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad \text{Equation 3-5}$$

If both sides of the Equation 3-5 are divided by γ_0 , we can obtain the autocorrelation function satisfying the same form as shown in Equation 3-6.

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad \text{Equation 3-6}$$

If $k = 1, 2, 3, \dots, p$ in Equation 3-6, we finally obtain Yule-Walker equations as in Equation 3-7.

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ \vdots & \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned} \quad \text{Equation 3-7}$$

Thus, the AR coefficients ($\phi_1, \phi_2, \dots, \phi_p$) can be estimated by autocorrelations ρ_k , and it can be expressed in the matrix form as $\Omega_p = \Phi P_p$, where

$$\Phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} \quad \Omega_p = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} \quad P_p = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & 1 \end{bmatrix}$$

3.1.3.2 Least Mean Square (LMS) Error Method for AR Coefficients Estimation

The LMS error method is an adaptive approach that updates the AR coefficients based on the feedback of prediction error [Adas98]. The i^{th} AR coefficient for next step Z_{t+1} ' prediction is denoted as ϕ_i^{t+1} , and they can be calculated based on Equation 3-8, where η is a constant called step size, and e_t is the prediction error between the actual value and the predicted value at time t .

$$\phi_i^{t+1} = \phi_i^t + \eta e_t Z_{t-i}' \quad \text{Equation 3-8}$$

The structure of LMS-based predictor can be illustrated by Fig 3-3. The predictor updates the AR coefficients based on the prediction error from the previous round of prediction, and uses the updated coefficients to estimate the next value. The predictor can start with any AR coefficients for the first round prediction. The coefficients will be adjusted based on Equation 3-8 in the following predictions. If the process is stationary, the coefficients will finally converge, and the prediction results will be the optimal with regard to the least mean square error.

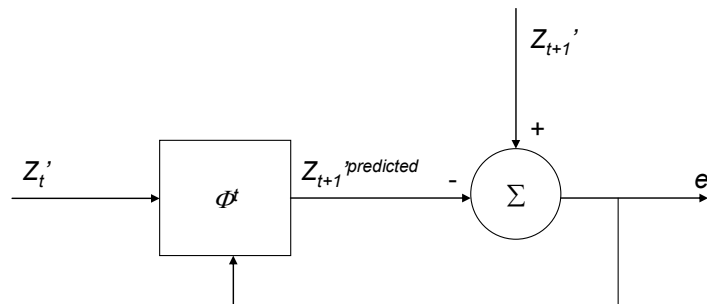


Fig 3-3: LMS-based predictor

The convergence speed is affected by the step size η . Using large η results in fast convergence and quick response to the changes. However, after convergence, the coefficients will have large fluctuations. Conversely, using small η results in slow convergence but less fluctuation after convergence.

The normalized LMS is a variation to the LMS, where the coefficients update is changed into Equation 3-9, where Φ_t is the vector of coefficients at time t , and L_d^t is the vector of actual value at time t used in the prediction, and $\|L_d^t\|^2 = L_d^t \bullet (L_d^t)^T$. The advantage of normalized LMS is that it is less sensitive to η , and normally the η is set to a value between 0 and 2 [LN02].

$$\Phi^{t+1} = \Phi^t + \frac{\eta e_t L_d}{\|L_d\|^2} \quad \text{Equation 3-9}$$

The LMS or normalized LMS predictor has been widely used in traffic prediction. For example, in FRR, the normalized LMS is used as the prediction approach to estimate the burst length. In ref [Adas98], an LMS-based predictor is used to forecast the frame information in the video traces. The LMS-based predictor is an adaptive AR model without requiring any prior knowledge of the traffic statistics or stationary assumption.

3.2 Challenges and Research Focuses of this Thesis

This thesis mainly focuses on exploring pro-active resource reservation mechanism, called the pre-booking mechanism, in the classical OBS, the hybrid network with both WRON and classical OBS operations, and the OBS with end-to-end connectivity acknowledgement protocols, such as the WR-OBS networks. The pro-active resource reservation mechanisms employ traffic prediction to achieve performance gains.

More specifically, in classical OBS, this research extends the FRR approach and proposes various forms of intelligence and a QoS provisioning method that does not have to penalise the low CoS traffic.

Secondly, this work explores average load prediction and proposes a mixed architecture that reserves and updates the wavelength routed lightpaths based on the predictions made, whilst, still delivering un-predicted traffic using classical OBS services. Additionally, this work also proposes a fast RWA mechanism based on ILP techniques. The proposed RWA mechanism can also be applied in current semi-permanent lightpath planning, such as to achieve the optimal or near-optimal placement with shortened solving time.

Finally, this thesis studies the pro-active reservation mechanism in OBS networks with end-to-end connectivity acknowledgement protocols, such as the WR-OBS networks. The work also investigates the performance when prediction error arises.

Chapter 4 Pre-booking Mechanism for Classical OBS Networks

This chapter describes a pro-active resource reservation mechanism applied to classical OBS networks. The pro-active resource reservation is realised by the newly proposed pre-booking mechanism, where traffic prediction technology is employed to forecast future traffic demands, and network resources are pro-actively reserved for those demands. The mechanism introduces edge intelligence and negotiation procedures, whilst, resource configuration is achieved using a classical OBS signalling protocol, such as the JET or JIT.

The chapter firstly describes the proposed network architecture, followed by the implementation and the evaluation of a simple pre-booking scenario without any advanced edge intelligence. It then moves to a more intelligent scenario in which traffic is allowed to be delayed at the edge node such that the overall blocking probability is reduced. Finally, a summary is given of the proposed scheme.

4.1 Proposed Network Architecture

The network architecture, as seen in Fig 4-1, mainly consists of Customers, Service Providers (SPs), and a Carrier Network Operator (CNO).

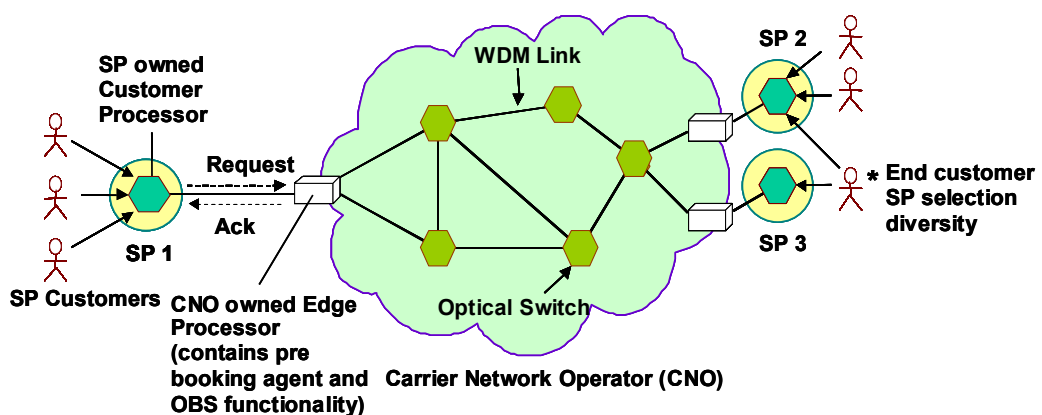


Fig 4-1: Proposed Architecture for Classical OBS

Customers are subscribers of SPs. They buy or explicitly pre-book network services or network resources from SPs.

SPs act as brokers between customers and the CNO. Via their Customer Processor (CP) function, SPs are equipped with mechanisms for traffic prediction, burst manipulation, QoS selection, and admission control, so that they can effectively interact with the core network management, in the guise of the Edge Processor.

The core network management operates as a multi-service commodity market, allowing various SPs to pre-book network resources at specified times and durations in the future. The core network is regarded as a series of “conveyer belts” between the network edges supporting transportation reservations. The management functionality needed to maintain this arrangement remains solely the responsibility of the Carrier Network Operator (CNO).

More specifically, the CNO owns the Edge Processor (EP) allocated at network edge as well as the core optical switches. The EP is comprised of Pre-booking management section and the OBS section. The Pre-booking section uses intelligence to select possible paths based on the information from the SP and the anticipated cost from the per-link cost gradient data structure; it employs signalling with the core to obtain updated cost information and to make bookings. It also communicates with the SP receiving transportation requests, and returns acknowledgments back to SP, possibly leading to further negotiation.

The OBS section in the EP functions as the normal OBS edge router. All the complexity is pushed to the edge of the CNO, mainly in the pre-booking section.

The core optical switches only maintain simple per-link reservation tables and rules for recording resource reservations.

4.1.1 SP’s Typical Operations

In each SP system, the Customer Processor (CP), typically located in a router device, predicts future traffic flows from its customer collector network(s) based on current and historical traffic records, and places the predicted traffic information into the resource reservation queue¹⁹.

Reservations can then be aggregated into bulk units and used to formulate appropriate resource

¹⁹ To provide differentiated services, separate reservation queues corresponding to several classes of service could be employed.

request messages to the CNO. It participates in admission negotiation with the EP owned by CNO. The SP's resource reservation signalling is able to negotiate with the CNO for variable length slots between specified network edge points based on their anticipated traffic patterns.

Traffic prediction is used to allow the SP to anticipate future traffic demands. If some of the traffic can be precisely forecasted long before it actually arises, the benefit of traffic prediction is not only more precise resource planning, but it also to reduce the need for on-the-fly offset-time calculations. This advanced planning also allows sufficient time to run optimisation algorithms and to permit negotiations with the CNO. More specifically, with the aid of traffic prediction, three classes of traffic can be identified:

The first category is *explicitly booked traffic*, which refers to the traffic transportation explicitly ordered by the SP's customer. For example, it can be a data transfer from client to a remote server, where the nature of the service and the start and stop times are known. The SP knows exactly what is required and can act as an agent for the customer to purchase the necessary CNO resources.

The second category of traffic is *implicitly predicted traffic*. This is the traffic that is anticipated by SP's inference intelligence based on historical records. The SP will typically book a base level of resources off the CNO over an extended period of time, suitable to accommodate a series of bursts. Depending upon the amount of risk the SP is prepared to take, the amount of CNO resources reserved could allow for more or less overbooking²⁰.

The third category is *transient traffic*, which arises without prior notice or cannot be predicted.

4.1.2 CNO's Operations

The Edge Processors owned by the CNO handle the three categories of traffic with different strategies. It is proposed that through estimation of path, cost signalling, and negotiation, the network resource can be better allocated.

²⁰ If the SP was to miscalculate the demand, it could then request an instantaneous booking of core resources from the CNO at the current time plus a small delta. However, given the immediacy of this request, the CNO would charge a premium to accept this request.

4.1.2.1 Pre-booking for Expected Future Traffic

For explicitly planned traffic and implicitly predicted traffic, the reservation request arrives far before the traffic arrives, so the system employs a set of pre-booking signalling and negotiation messages, as shown in Fig 4-2, to perform the carrier network reservation.

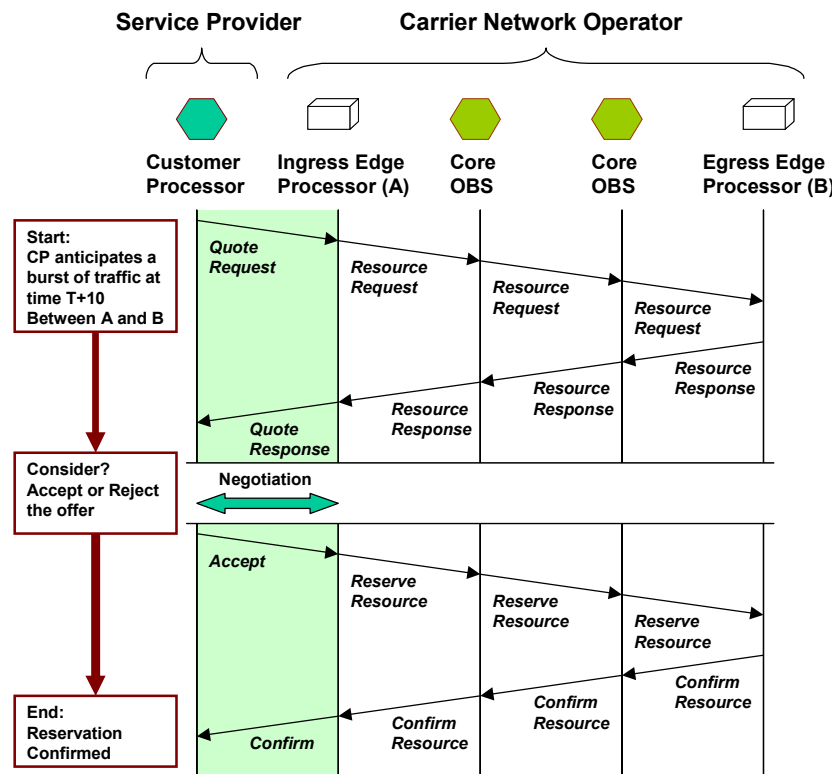


Fig 4-2: An Example – Negotiation and Signalling

The SP resource booking operation commences with SP determining the time and magnitude of the transport resources it will need between core network edge points, based on data held in the CP resource reservation queue. From this information the SP's CP first sends a quote request to the EP, including source/destination, burst length, proposed event time, selection criteria²¹, and "willingness to pay" information.

Upon receiving the quote request, the CNO's EP will start to calculate possible pathways for each request based on a link state database. The link state database maintained by EP comprises topology information along with link cost gradient details as shown in Fig 4-3. The per-link cost

²¹ Selection criteria are based on the Service Level Agreement (SLA). It can be end-to-end delay, or any other metrics.

gradients are based on known resource bookings and historical knowledge, such as busy hour trends. They are periodically updated to ensure that the information is not obsolete, and are mainly used to estimate the cost of the selected path at the requested future time instant, so that the path is more likely to meet SP's price requirement.

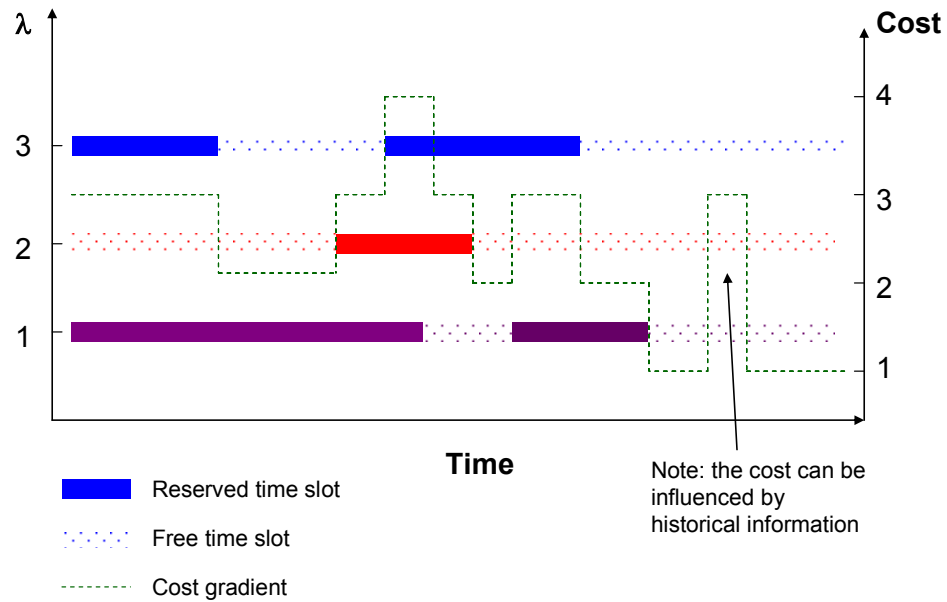


Fig 4-3: Cost Gradient on a Given Link

After the possible paths calculation, the EP then issues resource request message(s) along the calculated pathways to determine the true cost of the particular path at the allotted time. The cost information is accumulated in a hop-by-hop basis as the message traverses the core network. When a particular core switch receives a resource request message it looks through its resource reservation table for the intended next-hop link (shown in Fig 4-4) to determine its commitments. The resource reservation table in the OBS is a purely rule based table of known commitments²², containing the times that wavelengths are reserved or free for each link. The cost is obtained by taking a slice through the table from T_{start} to T_{stop} . The extent of these commitments determines the price, which is then added to the resource request message before it is forwarded to the next node along the path. If there are no resources available, the negative resource response message will be generated and returned to the originating EP.

²² There is no prediction or any other complexity associated with the table.

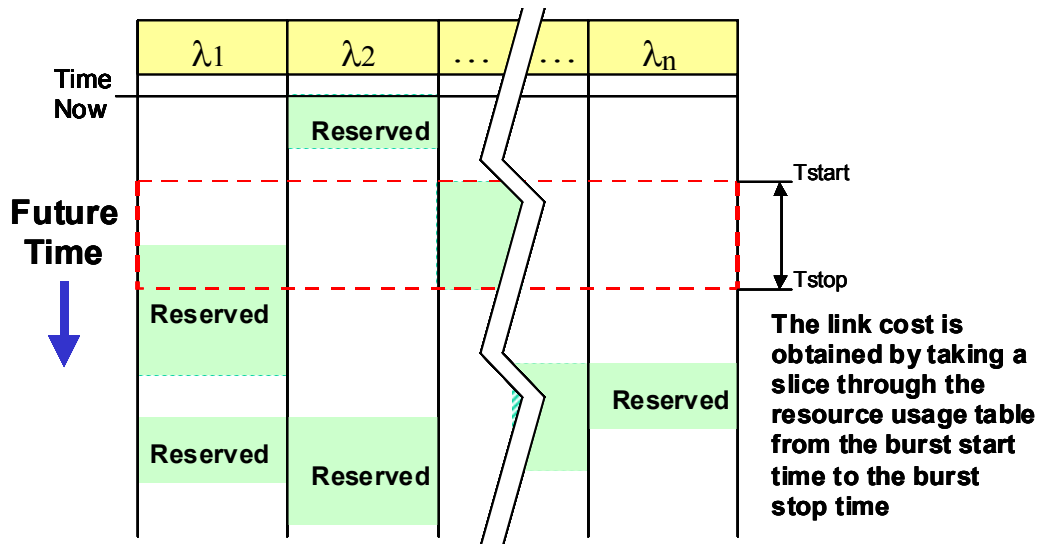


Fig 4-4: Example Resource Reservation Table in the Core Optical Switch

Assuming the resource request message manages to reach the core network egress point, it is converted into a resource response message and returned to the originating EP. This resource information is then used by the EP to construct a quote, and sent as a quote response message to the SP's CP. The SP then has to make a decision in a limited time to accept or reject the offer²³. It either replies with an accept message to ask for reservation confirmation, or does nothing if it wishes to reject the offer. If SP doesn't reply within the limited time, the offer is withdrawn. If the offer is accepted, a new signalling phase will be launched, which can be based on any OBS protocol such as JET or JIT, to finally reserve the resource²⁴. Otherwise, more negotiation will take place between the CP and the EP.

4.1.2.2 Normal OBS Reservation for Transient Traffic

Transient traffic arises without any notification and cannot be predicted, so there is little time for path estimation and further signalling or negotiation. In the design, it is handled using a traditional OBS resource allocation protocol, like JET or JIT.

In order to alleviate the network contention, the OBS routing table for transient traffic can be

²³ During this period, the price is guaranteed by the CNO.

²⁴ The final resource reservation can also be based on the two-way acknowledged signalling, which is shown in Fig 4-2.

dynamically updated based on the extent of commitments on each link²⁵.

4.2 Simple Pre-booking Scenario

The proposed pre-booking mechanism facilitates intelligent path selection, billing, and negotiation for predicted future traffic. However, in order to demonstrate the direct effect of pre-booking taking advantage of the time gained through traffic prediction, the simple pre-booking scenario skips any advanced intelligence, such as the path estimation and price adjustment functions. The current focus is to determine whether with 100% accurate traffic prediction and simple pre-booking, is it possible to reduce the blocking probability and increase the link utilisation. We attempt to answer questions such as: To what degree is the system able to improve the performance? How do different ratios of traffic mixture affect the results? Are there any drawbacks?

In the simple pre-booking scenario, it is assumed that (1) total traffic is grouped into predicted and non-predicted types. The predicted traffic refers to the future traffic that has been predicted or explicitly booked with 100 percent certainty²⁶, and the non-predicted traffic corresponds to the transient traffic that does not give any notification in advance or is not predicted; (2) currently, the differentiation between the explicitly subscribed traffic and the implicitly predicted traffic is ignored; (3) the buffer size at the edge components is assumed to be infinite; (4) no advanced intelligence and signalling, such as the path estimation function, the price confirmation, and negotiation, are involved in this model.

The predicted traffic reserves resources long before the actual traffic appears. What's more, as the path estimation and price collection procedures are ignored, the resource reservation for predicted traffic only involves the reservation signalling based on JET or JIT protocol. When non-predicted traffic arrives, the SP instantly asks for resources from the core using standard OBS signalling²⁷.

²⁵ This can be realized by a dynamic routing protocol running in background whereby link weights are affected by the level of resource commitments.

²⁶ The actual traffic prediction model is not the focus of this model.

²⁷ Current implementation takes JET as the underlying signalling protocol.

4.2.1 Simulation Model

The simulation was carried out using OPNET™ Modeller. In the simulation design, there is primarily an Edge Node model and a Core Node model. To start with, the functions of prediction, negotiation, and other intelligence are not considered. Therefore, there is no SP model, and the functions of source generation are included in the edge node model.

4.2.1.1 The Edge Node Model

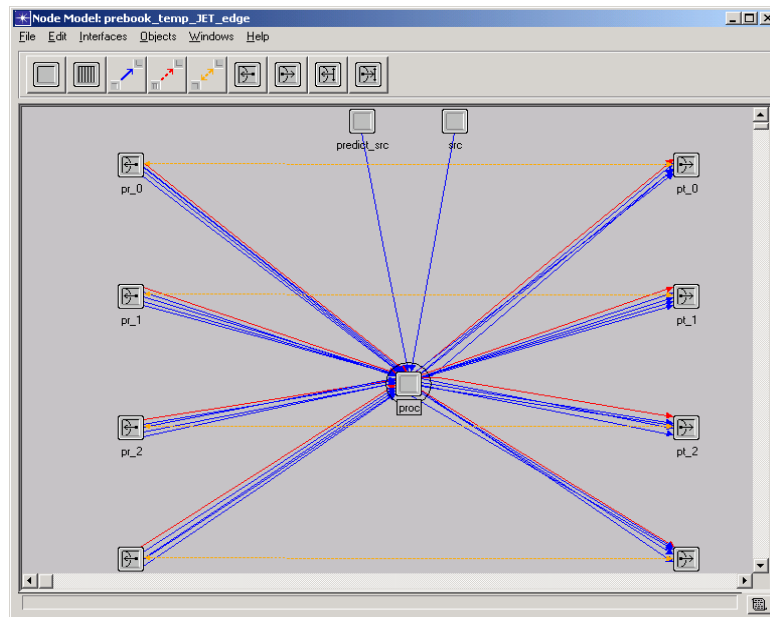


Fig 4-5: Edge Node Model

The edge node model represents the edge processor (EP) of the Carrier Network²⁸. The main functions of the current version include the traffic generation, BHP initiation, resource reservation, and burst delivery. The edge node uses a static routing table for both the predicted and non-predicted traffic rather than running a path estimation algorithm for the predicted traffic; it assumes infinite buffer capacity and it ignores the burst assembly mechanism, so the traffic generator directly generates bursts instead of packets.

Fig 4-5 shows the main processes in the edge node model. The two traffic generation processors are called “*predict_src*” and “*src*”²⁹. “*predict_src*” acts as a source predictor. Although a

²⁸ Currently, it contains two traffic generation processors, which should actually be in the SP node, to generate predicted and non-predicted traffic, respectively.

²⁹ Each generator is not a single customer. Instead, it represents the combined traffic coming into the node.

preliminary prediction engine is embedded in this stage, the “*predict_src*” periodically produces a batch of the predicted burst traffic³⁰ information according to a certain statistical distribution function (i.e. Poisson in this experiment). The predicted traffic information is then immediately sent to the “*proc*” processor asking for resource reservation. Unlike the “*predict_src*”, the “*src*” processor is a real traffic generator. It generates bursts one by one as the simulation progresses, and the burst will also be sent to “*proc*” asking for delivery.

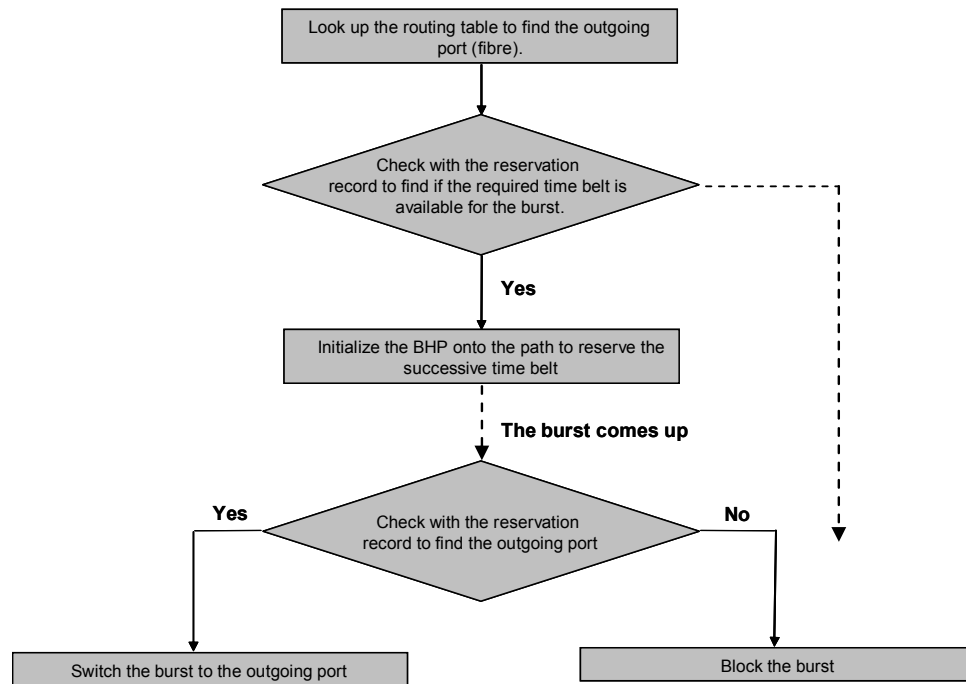


Fig 4-6: Handling the Predicted Traffic

The “*proc*” processor is the kernel of the edge processor. It treats predicted and non-predicted traffic slightly differently, as demonstrated in Fig 4-6 and Fig 4-7.

The difference is, for predicted traffic, it checks with the reservation record to see if the scheduled future time slot is available for the anticipated traffic using the LAUC-VF algorithm; while for the non-predicted traffic, it has to first calculate the offset-time and delay the burst in the queue such that there is enough time to issue the BHP onto the path to do the reservation also using the LAUC-VF algorithm.

³⁰ It is assumed that the traffic is anticipated by the predictor with 100 percent accuracy.

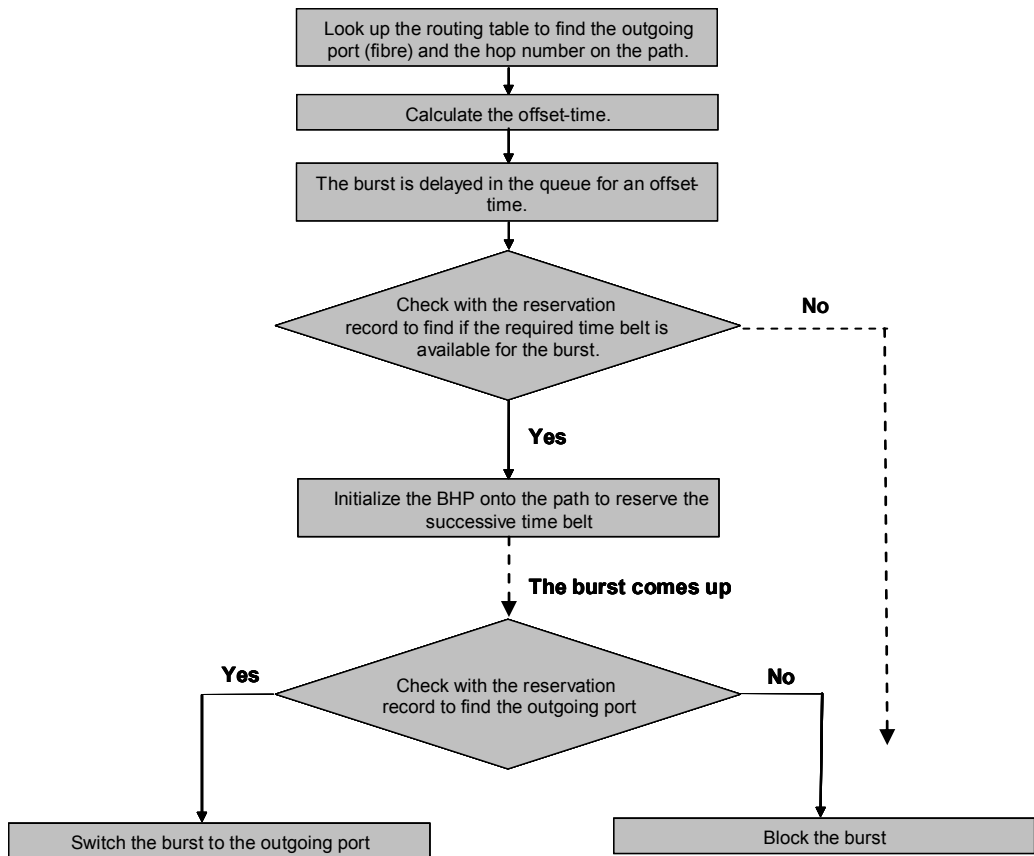


Fig 4-7: Handling the non-predicted traffic

4.2.1.2 The Core Node Model

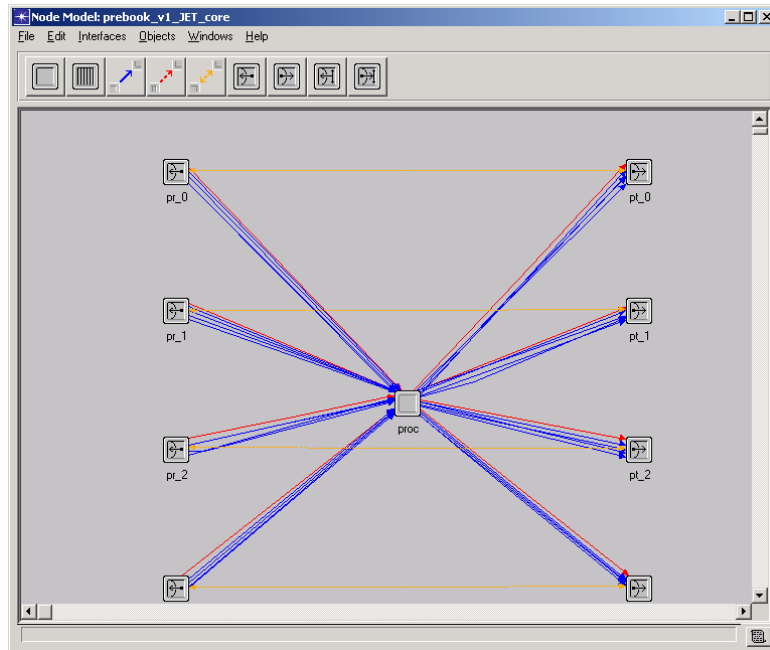


Fig 4-8: Core Node Model

The core node model represents a normal core OBS switch. The model itself seems very similar

to the edge node model, as shown in Fig 4-8.

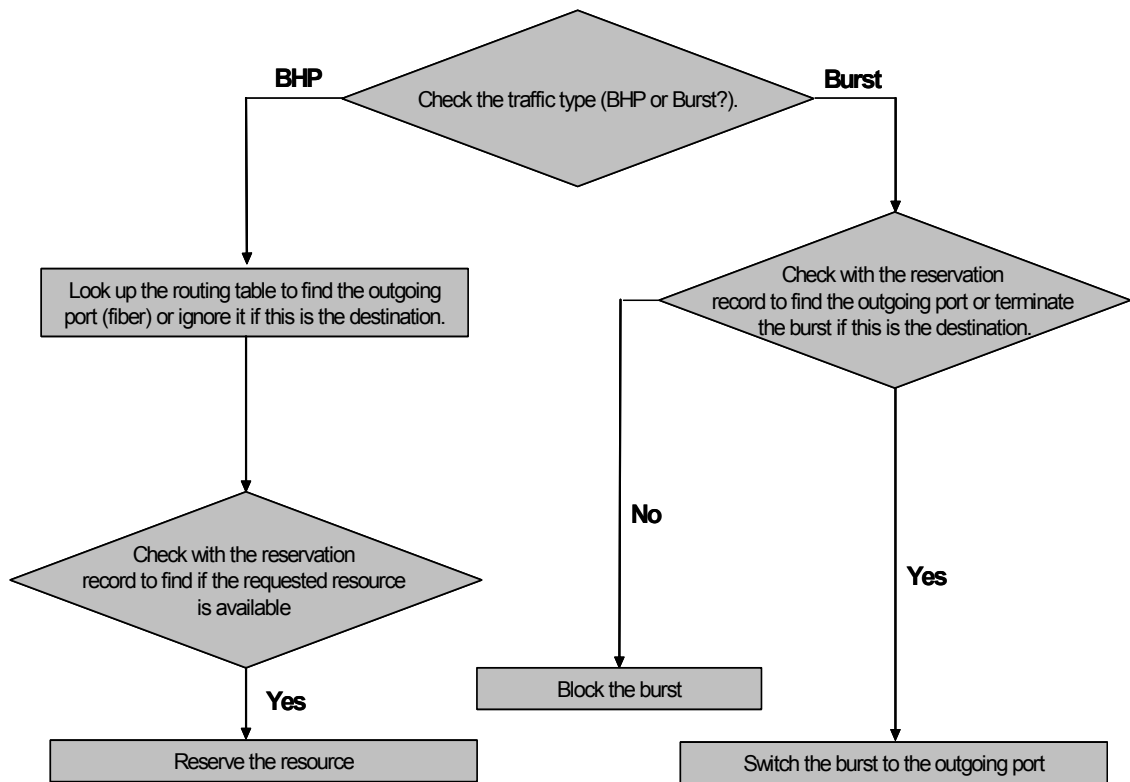


Fig 4-9: Workflow in the Core Node Model

In this model, there is no difference between the predicted and non-predicted traffic handling. The workflow inside the “*proc*” is shown in Fig 4-9. More specifically, the “*proc*” has two significant components, one for BHP processing, and the other for handling bursts.

Basically, if the “*proc*” receives BHP, based on the information carried by the BHP, it will look up a reservation table to see if the requested time slot is available on the outgoing link using the LAUC-VF algorithm, and then arrange the resource reservation for the upcoming burst. On the other hand, if the “*proc*” receives a burst, it will switch the burst if the reservation has been prepared for the burst, or block the burst if the reservation was not done in advance.

4.2.2 Validation of Simulation Model

As there are four key milestones during the modelling, the validation was mainly carried out to check the correctness of the important sub-models. The four milestone models are (1) the preliminary model that only involves packet transmission and receiving; (2) the LAUC-VF algorithm that is used to do the resource reservation; (3) the JET model that integrates standard

JET signalling together with the resource reservation model using LAUC-VF; (4) a simple pre-booking model which is built upon the standard JET model by introducing two groups of traffic at the source.

More details and snapshots for the validation are presented in appendix E.

4.2.3 Performance Results and Evaluation

The analysis model for the simple pre-booking mechanism is presented in appendix D. Both analysis and simulation results have been collected to investigate the blocking probability for predicted, non-predicted, and overall traffic using a simple pre-booking mechanism. The simulation is based on a simple two-node - one fibre arrangement. There are 4 wavelengths in the fibre, each wavelength operating at 10Mbps; the burst length is distributed by exponential process with mean 50k bytes; and the arrival process for burst data is Poissonian.

Fig 4-10 shows the blocking probability (averaged from 10 independent experiments, and presented with 95% confidence interval), where the ratio of predicted traffic to total traffic load is 50%, and the offered load refers to the ratio of average incoming bit rate to the wavelength rate.

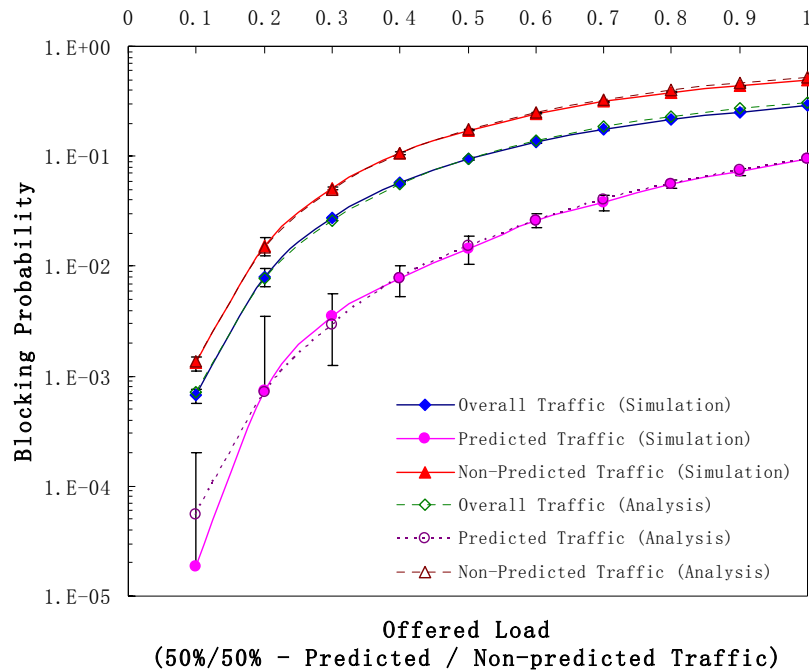


Fig 4-10: Simple Pre-booking Performance with a Constant Traffic Mixture

It can be seen that with increasing traffic load, the blocking probability of non-predicted traffic increases much more rapidly than predicted traffic. This is because the wavelength resources are reserved in advance by predicted traffic, which in turn blocks non-predicted traffic when resource contention arises. The simulation results closely match the analytical results, indicating the correctness of the simulation model.

The results shown in Fig 4-10 only consider resource pre-emption of the predicted traffic when the ratio of predicted traffic is kept at 50 percent to the total traffic. In order to investigate the effect of prediction with different ratios to the total traffic, another experiment was set up. Fig 4-11 investigates the blocking probability with different mixtures of predicted traffic and non-predicted traffic, whilst keeping the total traffic load at 0.8.

The result shows that by keeping total traffic load at 0.8 while increasing the ratio of predicted traffic from 0% to 100% of the total traffic, the blocking probability of the predicted traffic is always lower than that of the overall traffic, and the non-predicted traffic is always higher than that of the overall traffic. Meanwhile, the overall blocking probability remains constant as expected from Equation D-4.

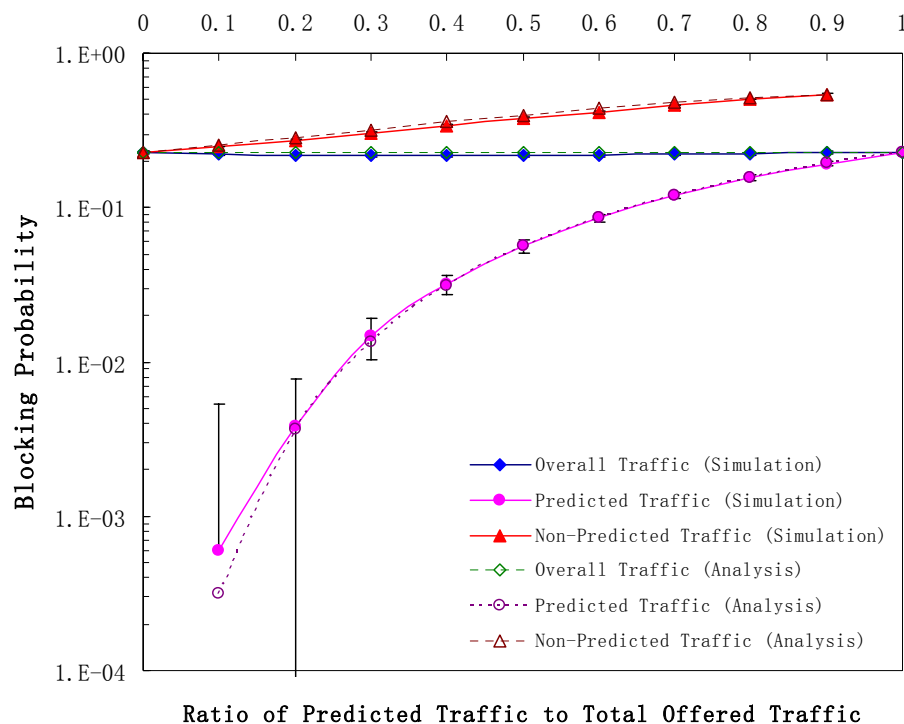


Fig 4-11: Simple Pre-booking Performance with Various Traffic Mixtures

Hence, the predicted traffic always yields better performance than the non-predicted traffic. However, although the pre-booking is able to bring benefits to predicted traffic, it degrades the performance of non-predicted traffic. Therefore, the simple traffic pre-booking cannot improve the overall traffic performance. However, it is expected that this situation can be improved by the introduction of more flexible pre-booking, where intelligence at the edge of the carrier network is able to offset burst arrivals to some degree.

4.3 Shifting Mechanism Evaluation

The results from the simple pre-booking system demonstrate that a pure pre-booking system cannot bring any benefits to the overall traffic performance. To further make advantage of the prediction, more intelligence needs to be deployed to proactively adjust the resource arrangements. The use of time-shifting is now considered with the purpose of enhancing the overall traffic performance.

The shift idea is rather simple: to shift the reservation time to later value(s) if the requested time and duration clashes with an existing reservation. Although the simulation model assumes an infinite buffer, there are in fact two ways to approximate this. Firstly, in the electrical domain large buffers could be used in the access network - the problem is the scale of the buffers required due to the large traffic volume. The second approach is to employ optical fibre loops. This enables discrete shifts, but it is inflexible. There may be a way to combine optical loops with electrical buffers. For example, when sufficient data is accumulated in the buffer, it is converted to light and sent around a loop. This frees up buffer space for a subsequent burst.

More importantly, as prediction typically allows for time to perform negotiations, whilst the non-predicted traffic has no such flexibility, the shift mechanism is only applied to the predicted traffic. Therefore, for predicted traffic, when the "Cost Request" signalling message is issued as a way of querying the availability of resources, the core switch may recommend a later reservation time if contention arises. The returned message will then be an offer with some delay offset from the original request. Through negotiation, the SP may accept or reject the revised offer.

4.3.1 Performance Results

Based on the simple pre-booking model, a shift method was added. Other simulation parameters and settings remain the same. The results in terms of blocking probability are collected.

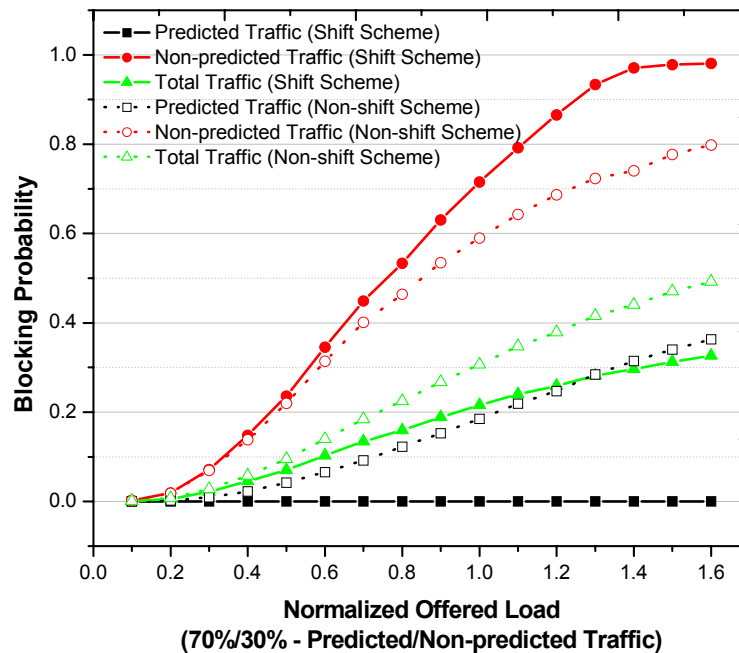


Fig 4-12: Shift Scheme Performance with a Constant Traffic Mixture

Fig 4-12 shows the blocking probability when the predicted traffic is 70 percent of the total traffic. The solid lines refer to the shift scheme, and the dashed lines refer to the simple pre-booking scheme. The blocking probability demonstrates that via shifting, the predicted traffic is able to achieve non-blocking all the way, and the overall blocking probability is lower than that of the simple pre-booking system. However, the line for the non-predicted traffic is very steep before it reaches 100 percent of blocking. This means that the shift has exploited resource pre-emption for predicted traffic, and therefore steals more resources from the non-predicted traffic. Although the non-predicted traffic is degraded, the overall performance is still better.

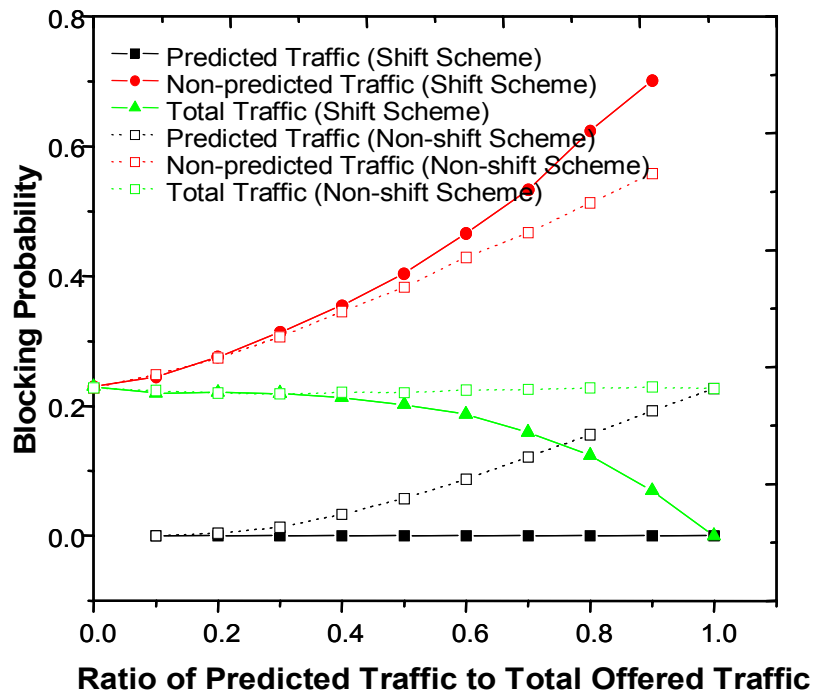


Fig 4-13: Shift Scheme Performance with Various Traffic Mixtures

Fig 4-13 investigates the scheme by varying the traffic proportions whilst keeping the overall load at 0.8. Similarly, the results show that no matter how the traffic is mixed, the predicted traffic is able to obtain 100 percent admission to the network resource. Although, a greater proportion of non-predicted traffic is blocked with the increase of predicted traffic, because the total amount of non-predicted traffic is decreasing, the overall performance still improves.

4.4 Summary

This part of the thesis proposes a novel pre-booking OBS network architecture to provide an economical and efficient optical network infrastructure for multi-service “bandwidth on demand” transport. It is proposed that by adding traffic prediction, intelligent resource provisioning, and negotiation at the network edge, the performance of a pure OBS system, such as blocking probability, can be improved.

Chapter 5 Pre-booking Mechanism for a Hybrid Network Architecture and the RWA Mechanism

5.1 Introduction

The previous chapter describes the application of a pro-active pre-booking mechanism in classical OBS networks, where the pre-booking relies on detailed burst length prediction. This chapter investigates the pre-booking mechanism when it is applied to large time scale average load prediction.

A hybrid network architecture which combines the Wavelength Routed Optical Network and the classical OBS is proposed. The core principle of the proposed scheme is that the network-planning components proactively reserve wavelength resources for implicitly predicted or explicitly pre-booked future traffic loads. This prior reservation concept inherits aspects of wavelength routing, providing end-to-end guaranteed lightpaths. The actual traffic will be principally delivered via these reserved lightpaths; however, if the demand exceeds the reserved lightpath capabilities, the excess traffic can be delivered via classical OBS signalling, but with its inherent risks. The architecture takes advantage of traffic load prediction to improve the wavelength routing efficiency, and the lightpath placement for predicted loads is dynamically adjusted based on updated traffic load information.

An RWA algorithm is useful to place the anticipated traffic flows appropriately across the network, and plays an important role in improving the network efficiency. Therefore, this chapter also studies various RWA algorithms based on Integer Linear Programming (ILP), and proposes a new two-stage RWA mechanism to obtain the optimal or near-optimal solution in a short time.

5.2 Proposed Network Architecture

The proposed network structure is shown in Fig 5-1. The WDM backbone is operated by a Carrier Network Operator (CNO), with various Service Providers (SPs) placed at the edge.

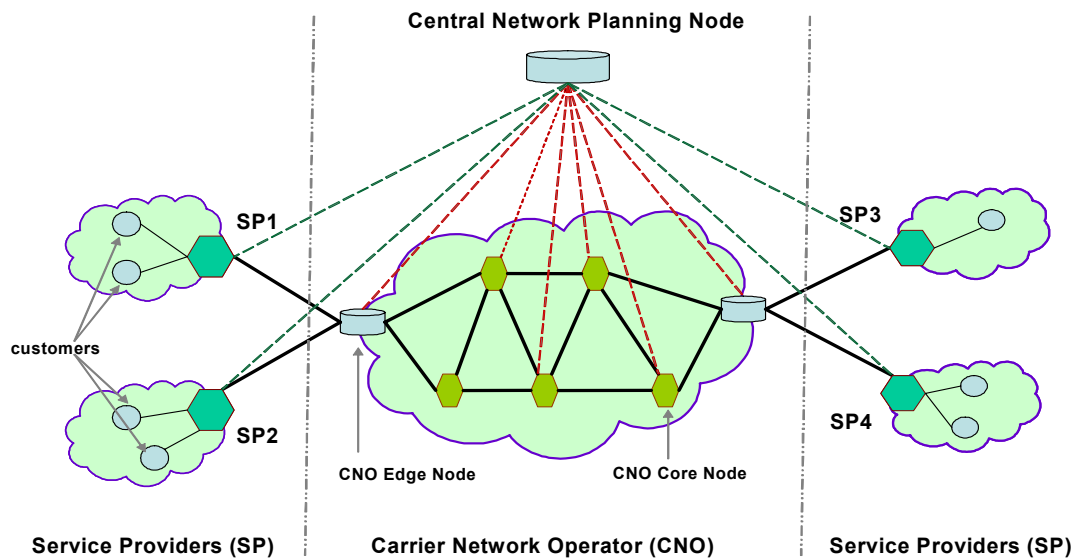


Fig 5-1: Proposed Architecture with Hybrid Operations

5.2.1 Service Providers (SPs)

Service Providers (SPs) manage customer traffic, and are responsible for making reservation requests either proactively, via the CNO central node, or reactively using the classical OBS signalling, by communicating directly with a CNO ingress optical switch. They are also responsible for burst assembly and QoS selection.

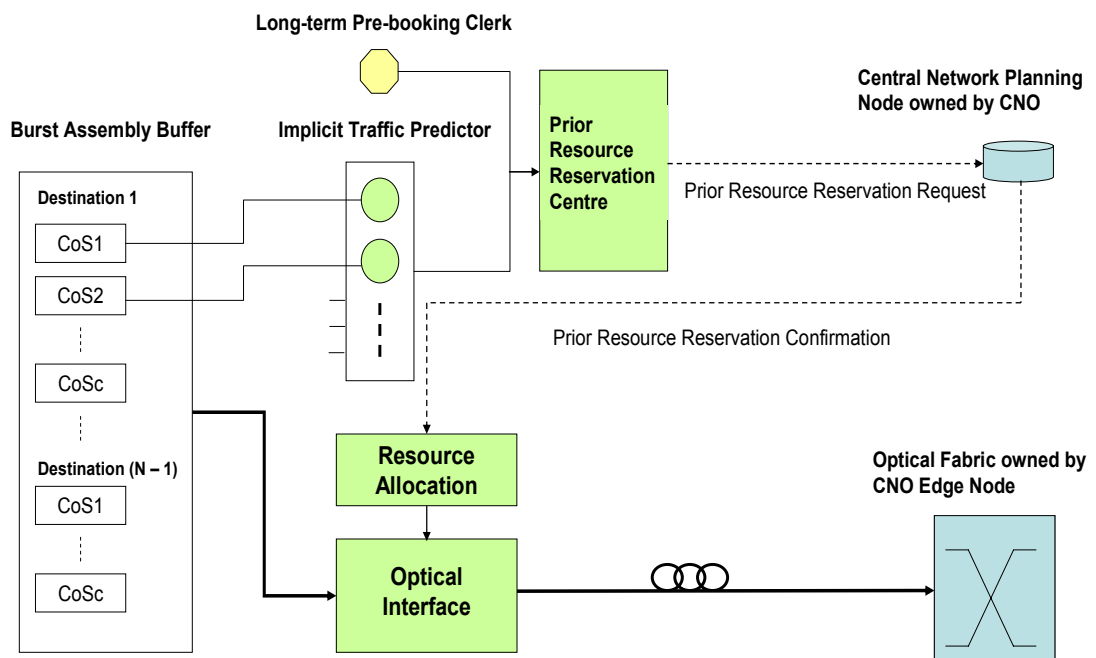


Fig 5-2: Functionality Blocks in a Service Provider

As shown in Fig 5-2, the prior resource reservation centre acts as a resource request agent,

liaising with the CNO's central node. It accepts explicit future pre-booking from customers via a long-term pre-booking clerk; it also predicts future end-to-end traffic demands based on historical traffic patterns and other implicit means available to it.

The pre-booked and predicted information is used to make decisions concerning resource reservations, as required. A specific lightpath reservation strategy must be employed to decide how many lightpaths should be reserved for the anticipated traffic. For instance, Fig 5-3 provides a step-shaped reservation strategy where the number of lightpaths is the ceiling value of the predicted load³¹. Such a strategy is able to alleviate the negative effects brought by the prediction error. For example, if the actual incoming bit rate is 5Gbps, whilst the anticipated incoming bit rate is 8Gbps, with each wavelength operating at 10Gbps, the number of light paths to be reserved for both actual load and anticipated load are 1.

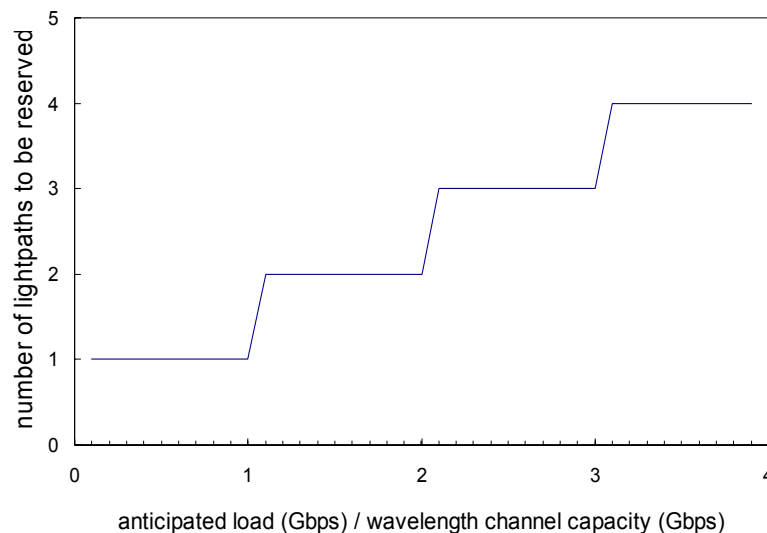


Fig 5-3: An Example of Resource Reservation Strategy

The SP is also responsible for sending traffic bursts into the CNO network at the agreed pre-booked time(s) that are then transported along the proactively reserved path. In order to increase the utilization of these pre-booked slots, multiple bursts may be “bulked” together and inserted into the same pre-booked transmission slot by delaying some of them a little. If the SP has reserved insufficient resources, it will typically arrange to employ classical OBS signalling

³¹ Load refers to the ratio of the average incoming bit rate to the wavelength channel capacity.

to deliver some bursts on-demand.

5.2.2 Carrier Network Operator (CNO)

As the CNO deals with both the prior reservation requests and classical OBS queries, it adopts a two-layer resource allocation strategy.

More specifically, this layered approach comprises a centralised proactive layer, dealing with pre-booking requests, and a distributed reactive lower-layer, providing basic OBS support at each optical switch. The proactive layer, focussed around the central network-planning node, maintains network state information, such as available resources, the network topology and a table of the pre-booked lightpath commitments, and performs the RWA for the pre-booking requests. The central node then sends signalling messages to the appropriate optical switches so that the necessary switching actions can be taken at the required time.

As to the reactive layer, each ingress node maintains a link state database whereby the cost of the links is related to the pre-booking commitments during the forthcoming interval. This information is used to perform an offline calculation of suitable path(s) from the given ingress point to the various egress nodes. This forms a source-routed routing table for OBS signalling. When a request to carry traffic arrives, i.e. which is not to be pre-booked, the Burst Head Packet (BHP) can be quickly loaded with the appropriate path vector from the routing table and sent out in an attempt to reserve the path. Periodically, based on future pre-booking commitments, the central network-planning node updates the ingress node link state databases for the next interval.

Such a pre-booking architecture provides a dynamically updated WR network with traditional OBS as a fallback solution. The prediction and pre-booking interval affects both the dynamicity and the scalability. If the interval is small (say burst by burst), the network is highly dynamic. However, it is very hard to predict detailed burst information with reasonable accuracy, and this also puts considerable stress on the central network planning node and the optical switches. Conversely, if the interval is too large, the architecture converges to static WR with a correspondingly low utilisation. Therefore, it is important to choose an appropriate pre-booking interval, such that the network resources can be quasi-dynamically provisioned, whilst alleviating the scalability concerns. According to the observed daily traffic patterns [JNM03], an

interval ranging from tens of minutes to several hours can be attempted, and choosing a suitable pre-booking interval is one of the topics for future work.

5.2.3 Potential Benefits

The proposed network architecture combines WRON and OBS in a single network environment, and takes advantage of the traffic inference to update the WRON lightpath placements. The benefits include the following two points.

Firstly, the combined architecture provides end-to-end guaranteed lightpaths and connectionless services simultaneously. This offers multiple choices for various services. For example, for latency sensitive and loss sensitive traffic, the WR lightpath is a good option. For other services that do not have strict requirements on delay and loss, the OBS can be the solution.

Secondly, the pre-booking mechanism helps to update the lightpath placement based on anticipated loads. This can avoid resource over-provisioning or resource exhaustion problems in the static WRON. Though the efficiency is affected by the prediction accuracy, the system can be improved by an appropriate resource subscription strategy, and the adjustment of resource subscription by comparing the real traffic load with the subscribed resource load.

5.3 RWA Algorithms

RWA is one of the most essential problems in WDM networks, and it determines paths and wavelengths for associated traffic to be delivered. In the proposed network architecture, RWA is required by the central planning node to produce the lightpath placement solution for the pre-booking requests.

Generally, due to the NP-complete nature of the problem, it is challenging to derive provably optimal or near-optimal solutions in a short time period for at least mid-sized wavelength constrained WDM networks. (For more background on RWA algorithms, please refer to chapter 2, section 2.2.2.)

The research here endeavours to create an improved ILP approach that is able to obtain optimal or near optimal solution within short time scale.

5.3.1 Problems under Study

The research in this thesis focuses on the Max-RWA problem in wavelength constrained WDM networks, which is one of the most common problems being considered [KS01][BCB04][JMT04][TMP02].

More specifically, the goal of Max-RWA is to achieve the maximum number of connections with the given number of wavelengths. In other words, the Max-RWA minimises the blocking probability for the traffic demand with the limited number of wavelengths. In practice, wavelength conversion is expensive, so most of the RWA algorithms should also impose the wavelength continuity constraint during the path selection procedure, such that each lightpath only consume one wavelength and no two link-joint lightpaths use the same wavelength. In this thesis, Max-RWA without wavelength conversion support is called as *wavelength constrained Max-RWA*, and *wavelength unconstrained Max-RWA*, otherwise.

5.3.2 Conventional ILP Models

There have been three major ILP models for the Max-RWA problem. They are path-based, flow-based, and source-based models, respectively [BCB04].

The path-based formulations consider all the possible loop-free paths within the model, and determine which path should be employed in the final optimal solution. This model leaves the path calculation to other external programmes, thus alleviating the difficulty of formulation. However, the number of the total paths increases exponentially with the network size, so the model has to compromise by taking only part of the paths in large-sized networks, thus losing the optimality guarantee [WD96].

The flow-based formulations take network topology links as variables, and the formulations determine which links should be employed for the corresponding flows. Although the number of variables in flow-based model is confined to the number of links, without the pre-knowledge of the paths, the model has to include more constraints to derive valid paths [JMT04].

The source-based formulations are similar to the flow-based formulations, but the model groups the connections with respect to their source nodes, and effectively reduces the number of variables and constraints [JMT04].

5.3.2.1 Path-based Formulations

The path-based formulations start with an enumeration of all the possible loop-free routes between all source-destination (sd) pairs, denoted as $R = \{rt_1, rt_2, \dots, rt_r\}$, assuming that the total number of the routes is r . Other parameters regarding the given network topology and the traffic demands are shown in Table 5-1.

Table 5-1: Parameters for Path-based ILP Formulations

Parameter	Size	Description
λ		Number of wavelengths on each unidirectional link
e		Number of unidirectional links in the network
sd		Number of source-destination pairs that yield at least one lightpath demands
D_p	sd	D_p is the number of lightpaths required to be setup for source-destination pair p
q_{ip}	$r \times sd$	$q_{ip} = 1$, if route rt_i is incident to source-destination pair p ; $q_{ip} = 0$, otherwise
b_{il}	$r \times e$	$b_{il} = 1$, if link l is part of route rt_i ; $b_{il} = 0$, otherwise

5.3.2.1.1 Wavelength Constrained Max-RWA

For wavelength constrained Max-RWA problem, the variables in path-based formulations are:

$$x_{iw} = \begin{cases} x_{iw} = 1, & \text{if route } rt_i \text{ and wavelength } w \text{ is employed as a lightpath} & (i = 1, 2, \dots, r) \\ x_{iw} = 0, & \text{otherwise} & (w = 1, 2, \dots, \lambda) \end{cases}$$

Objective: maximise the number of lightpaths being setup, as shown in Equation 5-1.

$$OBJ_{path-WC} = Max \left(\sum_{i=1}^r \sum_{w=1}^{\lambda} x_{iw} \right) \quad \text{Equation 5-1}$$

Constraint 1: the established lightpaths between source-destination pair, p , should not exceed the corresponding number of required lightpaths, D_p , as shown in Equation 5-2.

$$\sum_{i=1}^r \left(q_{ip} \times \left(\sum_{w=1}^{\lambda} x_{iw} \right) \right) \leq D_p \quad p = 1, 2, \dots, sd \quad \text{Equation 5-2}$$

Constraint 2: on each link, each wavelength should only be employed once, ensuring that the total number of flows travelled on each link does not exceed the number of wavelengths provided. The mathematical expression for this constraint is illustrated in Equation 5-3.

$$\sum_{i=1}^r (x_{iw} \times b_{il}) \leq 1 \quad w = 1, 2, \dots, \lambda \quad l = 1, 2, \dots, e \quad \text{Equation 5-3}$$

5.3.2.1.2 Wavelength Unconstrained Max-RWA

For the wavelength unconstrained Max-RWA problem, the lightpath can use any wavelengths along the route, so there is no need to distinguish wavelengths on the lightpath. The variables in this case are:

x_i , where x_i is the number of lightpaths being established using route rt_i , and it has to be an integer for $i = 1, 2, \dots, r$.

Objective: maximise the number of lightpaths being setup, as shown in Equation 5-4.

$$OBJ_{path-WUC} = Max \left(\sum_{i=1}^r x_i \right) \quad \text{Equation 5-4}$$

Constraint 1: the established lightpaths between source-destination pair, p , should not surpass the corresponding demands, D_p , as shown in Equation 5-5.

$$\sum_{i=1}^r (x_i \times q_{ip}) \leq D_p \quad p = 1, 2, \dots, sd \quad \text{Equation 5-5}$$

Constraint 2: the total number of flows travelled on each link should not exceed the wavelength number λ , as shown in Equation 5-6.

$$\sum_{i=1}^r (x_i \times b_{il}) \leq \lambda \quad l = 1, 2, \dots, e \quad \text{Equation 5-6}$$

5.3.2.2 Flow-based Formulations

In flow-based formulations, the network topology and the lightpath demand are the given

parameters. Specifically, each lightpath demand is described by its source and destination nodes, so the k^{th} lightpath demand is denoted as $lp_k=(s_k, d_k)$, where s_k is the source node and d_k is the destination node for this demand. Assuming that the total number of lightpaths required to be established is K , $LP=\{lp_1, lp_2, \dots, lp_K\}$ represents a vector containing all the lightpath demands that are required to be established. Other parameters are represented by the notations as in Table 5-2.

Table 5-2: Parameters for Flow-based ILP Formulations

Parameter	Description
λ	Number of wavelengths on each unidirectional link
e	Number of unidirectional links in the network
n	Number of nodes in the network
l	Link l
$V=\{v_1, v_2, \dots, v_n\}$	A vector of nodes in the network
$g^+(v_i)$	The set of outgoing links at node i
$g^-(v_i)$	The set of incoming links at node i

5.3.2.2.1 Wavelength Constrained Max-RWA

For wavelength constrained Max-RWA, there are two groups of binary variables in the flow-based formulations. The variables are:

$$x_k^w = \begin{cases} 1, & \text{if the } k^{th} \text{ lightpath demand is established by employing wavelength } w \\ 0, & \text{otherwise} \end{cases}$$

$$x_{kl}^w = \begin{cases} 1, & \text{if the } k^{th} \text{ lightpath demand is established by employing wavelength } w \text{ on link } l \\ 0, & \text{otherwise} \end{cases}$$

Objective: maximise the total number of established lightpaths, as shown in Equation 5-7.

$$OBJ_{flow-WC} = Max \left(\sum_{k=1}^K \sum_{w=1}^{\lambda} x_k^w \right) \quad \text{Equation 5-7}$$

Constraint 1: for each lightpath demand, the node v_i , excluding the corresponding source and destinations nodes, should yield the same number of incoming and outgoing flows using wavelength w , as shown in Equation 5-8.

$$\sum_{l \in g^+(v_i)} x_{kl}^w = \sum_{l \in g^-(v_i)} x_{kl}^w \quad k = 1, 2, \dots, K \quad w = 1, 2, \dots, \lambda \quad v_i \in V \setminus \{s_k, d_k\} \quad \text{Equation 5-8}$$

Constraint 2: for the k^{th} lightpath demand and its source node s_k , the total number of outgoing flows using the wavelength w should indicate whether the corresponding lightpath demand is established by employing wavelength w , as shown in Equation 5-9.

$$\sum_{l \in g^+(s_k)} x_{kl}^w = x_k^w \quad k = 1, 2, \dots, K \quad w = 1, 2, \dots, \lambda \quad \text{Equation 5-9}$$

Constraint 3: for the k^{th} lightpath demand and its destination node d_k , the total number of incoming flows using the wavelength w should indicate whether the corresponding lightpath demand is established by employing wavelength w , as shown in Equation 5-10.

$$\sum_{l \in g^-(d_k)} x_{kl}^w = x_k^w \quad k = 1, 2, \dots, K \quad w = 1, 2, \dots, \lambda \quad \text{Equation 5-10}$$

Constraint 4: for the k^{th} lightpath demand, the total number of incoming flows using the wavelength w at the source node s_k , and the number of outgoing flows using the wavelength w at the destination node d_k , have to be zero, as shown in Equation 5-11.

$$\sum_{l \in g^-(s_k)} x_{kl}^w = \sum_{l \in g^+(d_k)} x_{kl}^w = 0 \quad k = 1, 2, \dots, K \quad w = 1, 2, \dots, \lambda \quad \text{Equation 5-11}$$

Constraints 1, 2, 3 and 4 work together to assign a lightpath that yields a unique wavelength throughout for the lightpath demand k .

Constraint 5: the wavelength w on link l should be assigned to no more than 1 lightpaths, as in Equation 5-12.

$$\sum_{k=1}^K x_{kl}^w \leq 1 \quad l = 1, 2, \dots, e \quad w = 1, 2, \dots, \lambda \quad \text{Equation 5-12}$$

Constraint 6: there should be only one wavelength assigned for the k^{th} lightpath demand, if the k^{th} lightpath demand is established, as shown in Equation 5-13.

$$\sum_{w=1}^{\lambda} x_k^w \leq 1 \quad k = 1, 2, \dots, K \quad \text{Equation 5-13}$$

Constraint 7: the value of x_{kl}^w should be no more than that of x_k^w , as shown in Equation 5-14.

$$x_{kl}^w \leq x_k^w \quad k = 1, 2, \dots, K \quad w = 1, 2, \dots, \lambda \quad l = 1, 2, \dots, e \quad \text{Equation 5-14}$$

5.3.2.2.2 Wavelength Unconstrained Max-RWA

Without wavelength constraints, the formulations do not have to consider the wavelength, so the two groups of binary variables used are as follows:

$$x_k = \begin{cases} 1, & \text{if the } k^{\text{th}} \text{ lightpath demand is established} \\ 0, & \text{otherwise} \end{cases}$$

$$x_{kl} = \begin{cases} 1, & \text{if the } k^{\text{th}} \text{ lightpath demand is established by employing link } l \\ 0, & \text{otherwise} \end{cases}$$

Objective: maximise the total number of established lightpaths, as shown in Equation 5-15.

$$OBJ_{\text{flow-WUC}} = \text{Max} \left(\sum_{k=1}^K x_k \right) \quad \text{Equation 5-15}$$

Constraint 1: for each lightpath demand, the node v_i , excluding the corresponding source and destinations nodes, should yield the same number of incoming and outgoing flows, as shown in Equation 5-16.

$$\sum_{l \in g^+(v_k)} x_{kl} = \sum_{l \in g^-(v_k)} x_{kl} \quad k = 1, 2, \dots, K \quad v_i \in V \setminus \{s_k, d_k\} \quad \text{Equation 5-16}$$

Constraint 2: for the k^{th} lightpath demand and its source node s_k , the total number of outgoing flows should indicate whether the corresponding lightpath demand is established, as shown in Equation 5-17.

$$\sum_{l \in g^+(s_k)} x_{kl} = x_k \quad k = 1, 2, \dots, K \quad \text{Equation 5-17}$$

Constraint 3: for the k^{th} lightpath demand and its destination node d_k , the total number of

incoming flows should indicate whether the corresponding lightpath demand is established, as shown in Equation 5-18.

$$\sum_{l \in g^-(d_k)} x_{kl} = x_k \quad k = 1, 2, \dots, K \quad \text{Equation 5-18}$$

Constraint 4: for the k^{th} lightpath demand, the total number of incoming flows at the source node s_k , and the number of outgoing flows at the destination node d_k , should both be zero, as shown in Equation 5-19.

$$\sum_{l \in g^-(s_k)} x_{kl} = \sum_{l \in g^+(d_k)} x_{kl} = 0 \quad k = 1, 2, \dots, K \quad \text{Equation 5-19}$$

Constraints 1, 2, 3 and 4 work together to assign a lightpath for lightpath demand k .

Constraint 5: the total number of flows travelling through the link l should be no more than the number of wavelengths λ , as shown in Equation 5-20.

$$\sum_{k=1}^K x_{kl} \leq \lambda \quad l = 1, 2, \dots, e \quad \text{Equation 5-20}$$

Constraint 6: the value of x_k^l should be no more than that of x_k , as shown in Equation 5-21.

$$x_{kl} \leq x_k \quad k = 1, 2, \dots, K \quad l = 1, 2, \dots, e \quad \text{Equation 5-21}$$

5.3.2.3 Source-based Formulations

Source-based formulations use almost the same parameters as in the flow-based formulations, and it introduces new parameters D_s to represent the set of destination nodes for lightpath demands originating from the node v_s , T_{sd} to represent the number of lightpaths to be established between node v_s and v_d , and T_s to represent the number of required lightpaths from node v_s .

5.3.2.3.1 Wavelength Constrained Max-RWA

The source-based formulations for wavelength constrained Max-RWA problem employ the following variables to describe the problem:

$$x_{sl}^w = \begin{cases} 1, & \text{if a lightpath demand originating from node } v_s \text{ is established by employing} \\ & \text{wavelength } w \text{ on link } l \\ 0, & \text{otherwise} \end{cases}$$

Objective: maximise the number of established lightpath, as shown Equation 5-22.

$$OBJ_{source-WC} = Max \left(\sum_{(v_s, v_d) \in V: T_{sd} > 0} \left(\sum_{w=1}^{\lambda} \sum_{l \in g^-(v_d)} x_{sl}^w - \sum_{w=1}^{\lambda} \sum_{l \in g^+(v_d)} x_{sl}^w \right) \right) \quad \text{Equation 5-22}$$

Constraint 1: for all the lightpath demands originating from node v_s , the node v_i , which is exclusive from the nodes in D_s and the v_s itself, should have the same number of incoming and outgoing flows that employ wavelength w . This can be expressed in Equation 5-23.

$$\sum_{l \in g^+(v_i)} x_{sl}^w = \sum_{l \in g^-(v_i)} x_{sl}^w \quad w = 1, 2, \dots, \lambda \quad v_s \in V : T_s > 0 \quad v_i \in V \setminus (D_s \cup \{v_s\}) \quad \text{Equation 5-23}$$

Constraint 2: for lightpath demands originating from node v_s , the difference between the number of total incoming and outgoing flows at the destination node v_d is equal to the number of established lightpaths from node v_s to node v_d , which should be less than the demand number. It can be expressed as in Equation 5-24.

$$\sum_{w=1}^{\lambda} \sum_{l \in g^-(v_d)} x_{sl}^w - \sum_{w=1}^{\lambda} \sum_{l \in g^+(v_d)} x_{sl}^w = x_{sd} \leq T_{sd} \quad v_s \in V : T_s > 0, v_d \in D_s \quad \text{Equation 5-24}$$

Where x_{sd} is a temporary variable representing the number of established lightpaths from node v_s to node v_d .

Constraint 3: for lightpath demands originating from node v_s and ending at node v_d , the number of incoming flows using wavelength w should be no less than that of outgoing flows at the destination, as shown in Equation 5-25.

$$\sum_{l \in g^-(v_d)} x_{sl}^w \geq \sum_{l \in g^+(v_d)} x_{sl}^w \quad w = 1, 2, \dots, \lambda \quad v_s \in V : T_s > 0 \quad v_d \in D_s \quad \text{Equation 5-25}$$

The constraints 1, 2, and 3 act together to assign a unique wavelength for each lightpath.

Constraint 4: the wavelength w on link l should be at most employed once, as expressed in Equation 5-26.

$$\sum_{v_s \in V: T_s > 0} x_{sl}^w \leq 1 \quad w = 1, 2, \dots, \lambda \quad l = 1, 2, \dots, e \quad \text{Equation 5-26}$$

5.3.2.3.2 Wavelength Unconstrained Max-RWA

With wavelength conversion support, the problem can be formulated using the following

variables:

x_{sl} , where x_{sl} is number of established lightpaths demands originating from node v_s , on link l .

Objective: maximise the number of established lightpaths, as shown Equation 5-27.

$$OBJ_{source-WUC} = Max \left(\sum_{(v_s, v_d) \in V: T_{sd} > 0} \left(\sum_{l \in g^-(v_d)} x_{sl}^l - \sum_{l \in g^+(v_s)} x_{sl}^l \right) \right) \quad \text{Equation 5-27}$$

Constraint 1: for all the lightpath demands originating from node v_s , the node v_i , which is exclusively from the nodes in D_s and v_s itself, should have the same number of incoming and outgoing flows. This can be expressed as in Equation 5-28.

$$\sum_{l \in g^+(v_i)} x_{sl} = \sum_{l \in g^-(v_i)} x_{sl} \quad v_s \in V: T_s > 0 \quad v_i \in V \setminus (D_s \cup \{v_s\}) \quad \text{Equation 5-28}$$

Constraint 2: for lightpath demands originating from node v_s , the difference between the total number of incoming and outgoing flows at the destination node v_d is equal to the number of established lightpaths from node v_s to node v_d , which should not exceed the demanded number of lightpaths, as in Equation 5-29.

$$\sum_{l \in g^-(v_d)} x_{sl} - \sum_{l \in g^+(v_d)} x_{sl} = x_{sd} \leq T_{sd} \quad v_s \in V: T_s > 0, v_d \in D_s \quad \text{Equation 5-29}$$

Where x_{sd} is a temporary variable representing the number of established lightpaths from node v_s to node v_d .

Constraint 3: for lightpath demands originating from node v_s and ending at node v_d , the number of incoming flows should be no less than that of outgoing flows at the destination, as shown in Equation 5-30.

$$\sum_{l \in g^-(v_d)} x_{sl} \geq \sum_{l \in g^+(v_d)} x_{sl} \quad v_s \in V: T_s > 0, v_d \in D_s \quad \text{Equation 5-30}$$

The constraints 1, 2, and 3 act together to assign lightpaths.

Constraint 4: on link l , there should be at most λ flows travelling through, as expressed in Equation 5-31.

$$\sum_{v_s \in V: T_s > 0} x_{sl} \leq \lambda \quad l = 1, 2, \dots, e \quad \text{Equation 5-31}$$

5.3.2.4 Model Size Comparison

Path-based, flow-based, and source-based ILP models employ different variables and parameters to express the objective and constraint functions, but they yield the same upper bound value provided by the optimal solution of their linear programming relaxation³² [BCB04].

When the models are solved by the solver, the solver firstly relaxes the integer variables to real variables, to derive the optimal solution under linear programming relaxation. The solver then employs a branch and bound algorithm to search a valid integer solution within the space resulting from the previous process. The more integer variables and constraints, the longer the solver may take to produce the final solution. Therefore, adopting the model with the least number of variables and constraints can help to obtain the solution more quickly.

Table 5-3 compares the three models in terms of their number of variables and constraints, where WC refers to Wavelength Constrained Max-RWA, WUC corresponds to Wavelength Unconstrained Max-RWA problem.

Table 5-3: Number of Variables and Constraints for the Three Models

	Path-based		Flow-based		Source-based	
	WC	WUC	WC	WUC	WC	WUC
variables	$r \times \lambda$	r	$d \times \lambda \times (1+e)$	$d \times (1+e)$	$n \times \lambda \times e$	$n \times e$
constraints	$sd + \lambda \times e$	$sd + e$	$d \times \lambda \times (n+e+2) + e \times \lambda + d$	$d \times (n+e+2) + e$	$n(n-1)(\lambda+1) + e\lambda$	$2n(n-1) + e$

Notations:

- λ Number of wavelengths on each link
- n Number of nodes in the network
- e Number of unidirectional links in the network
- r Number of routes for the path-based model

³² Linear programming relaxation means that the model is solved by relaxing the integer variables to real number variables.

sd	Number of source-destination pairs that require at least one lightpath
d	Number of total lightpath demands

Table 5-3 illustrates that for all the three models, the number of variables for wavelength unconstrained Max-RWA problem is $1/\lambda$ that of wavelength constrained Max-RWA problem. Additionally, the number of constraints for wavelength unconstrained case is also approximately $1/\lambda$ that of wavelength constrained problem. Therefore, the model size for wavelength unconstrained Max-RWA problem will be smaller than that of wavelength constrained problem, which further indicates a shorter problem solving time than the wavelength constrained case. The size difference between the two problems becomes significant when the number of wavelengths gets large.

When the three models are compared with each other, the path-based model size depends on the number of total possible loop-free routes in the network, i.e. the value of r . However, this number increases exponentially with the number of network nodes and links. For example, in the US NSFnet with 14 nodes and 42 unidirectional links, the total number of loop-free routes are 13187, whilst, in the European EONnet with 20 nodes and 78 unidirectional links, the total number of loop-free routes is more than ten thousand, and the number of loop-free routes with at most 6 hops is as high as 12014. Therefore, the path-based model is not suitable for large networks unless only parts of the routes are considered in the model.

Regarding flow-based and source-based models, the number of variables and constraints are represented in Fig 5-4. Fig 5-4 provides an example where the number of network nodes is 14, i.e. $n=14$; the number of unidirectional links $e=42$; and the number of wavelengths on each link $\lambda=16$. The figure shows that the flow-based model size increases quickly with the value of d , which is the total number of lightpath demands in the traffic matrix. However, the number of variables and constraints for source-based model does not change with d . For instance, in the wavelength constrained case, if the total number of lightpath demands is less than 5, the flow-based model requires less variables and constraints than of the source-based model, whilst, the flow-based model size becomes significantly larger than that of source-based model when d exceeds 15. Therefore, in a mid-sized network with tens of nodes and links, if the total number of lightpath demands is not too small, the source-based model will yield the least number of

variables and constraints, which should result in a shorter solution time.

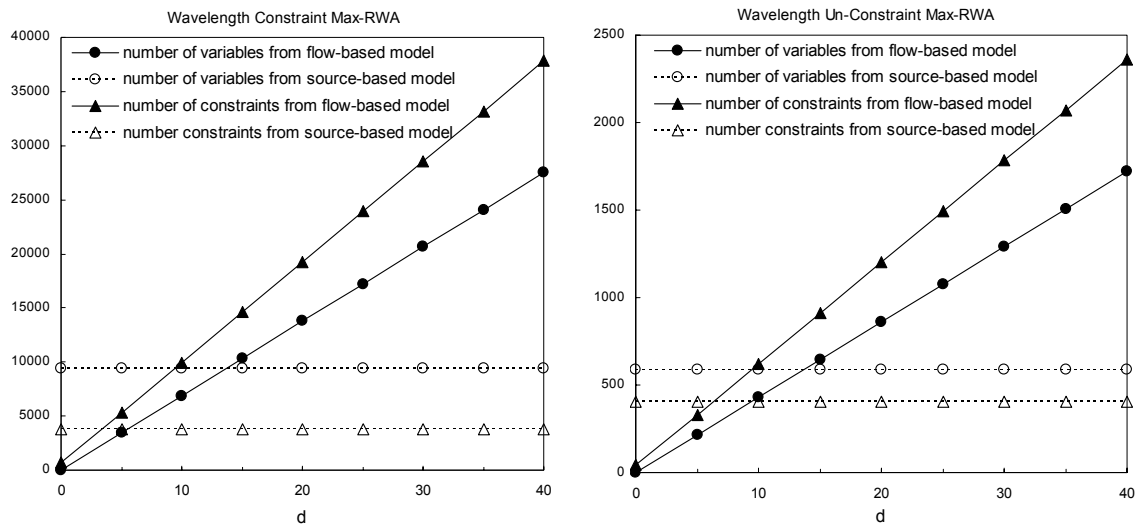


Fig 5-4: Comparison of Model Size between the Flow-based and the Source-based Models

5.3.3 Proposed Approach for Wavelength Constrained Max-RWA Problem

Though the wavelength constrained Max-RWA problem can be formulated into ILP models via the three conventional approaches, experiments using the lp-solve³³ show that these direct ILP models cannot produce any solutions even when they were ran for several days with popularly used networks³⁴ and traffic matrices³⁵.

Nevertheless, the corresponding ILP models for the wavelength unconstrained Max-RWA problem yield less variables and constraints, so they are more readily tractable than the ones with wavelength continuity constraints (Experimental results regarding the objective value and the execution time of the three models will be presented in a later section).

Therefore, this research proposes to decompose the wavelength constrained Max-RWA problem

³³ lp_solve is a free linear (integer) programming solver based on the revised simplex method and the Branch-and-bound method for integers. The software can be downloaded from <http://lpsolve.sourceforge.net>.

³⁴ The 14-node US NSFnet, and the 20-node European EONnet.

³⁵ The matrices used in the experiments are the ones used by Krishnaswamy and Sivarajan in ref [KS01]. The matrix were used by many other researchers [BCB04].

into two stages. In the first stage, the problem is solved permitting the wavelength conversion to be temporally allowed. The route candidates and the upper bound is obtained from a wavelength unconstrained Max-RWA ILP model. In the second stage, graph colouring technology is employed to assign each path with a unique wavelength, and assuring no two link-joint lightpaths use the same wavelength.

As it can be proved that the optimal objective value derived from the first stage is an upper-bound for the final wavelength constrained Max-RWA problem, the goal of the second stage is to find a valid colouring such that all the connections derived from the first stage can be setup using the given number of wavelengths. If this goal is achieved, it can be confirmed that the optimal solution is found. However, if there are insufficient wavelengths to set up all the lightpaths derived from the first stage, the upper-bound value will be able to provide an idea of how good the solution is from the upper-bound.

5.3.3.1 The First-Stage ILP and the Upper-Bound

In the first stage, an ILP model for the wavelength unconstrained Max-RWA is adopted. As the source-based ILP model yields a small number of variables and constraints, it is selected as an example to demonstrate the mechanics of the proposed approach.

Importantly, the objective value derived from the wavelength unconstrained Max-RWA ILP is no less than that from the corresponding wavelength constrained Max-RWA ILP, as expressed in Equation 5-32.

$$Obj_{un-constraint} \geq Obj_{constraint} \quad \text{Equation 5-32}$$

Taking the source-based ILP models as an example, and assuming that the optimal solution for the wavelength constrained Max-RWA ILP is $S_{constraint} = \{x_{sl}^w\}$, then, if z_s^l corresponds to the number of established flows originating from the node v_s on link l , z_s^l can be expressed by x_{sl}^w as $z_s^l = \sum_{w=1}^{\lambda} x_{sl}^w$.

Since all the x_{sl}^w comply with Equation 5-23, the equation can be transformed to Equation 5-33, if both sides of the equation are summed up through the wavelength space.

$$\sum_{w=1}^{\lambda} \left(\sum_{l \in g^+(v_i)} x_{sl}^w \right) = \sum_{w=1}^{\lambda} \left(\sum_{l \in g^-(v_i)} x_{sl}^w \right) \Leftrightarrow \sum_{l \in g^+(v_i)} z_s^l = \sum_{l \in g^-(v_i)} z_s^l \quad \text{Equation 5-33}$$

where $v_s \in V : T_s > 0, v_i \in V \setminus (D_s \cup \{v_s\})$

Equation 5-24 can be directly changed into Equation 5-34 by using the relationship between x_{sl}^w and z_s^l .

$$\sum_{l \in g^-(v_d)} z_s^l - \sum_{l \in g^+(v_d)} z_s^l \leq T_{sd} \quad v_s \in V : T_s > 0, v_d \in D_s \quad \text{Equation 5-34}$$

Similarly, the Equation 5-25 and Equation 5-26 turn into Equation 5-35 and Equation 5-36 by summing up both sides through the wavelength space.

$$\sum_{l \in g^-(v_d)} z_s^l \geq \sum_{l \in g^+(v_d)} z_s^l \quad v_s \in V : T_s > 0, v_d \in D_s \quad \text{Equation 5-35}$$

$$\sum_{v_s \in V : T_s > 0} z_s^l \leq \lambda \quad l \in L \quad \text{Equation 5-36}$$

Equation 5-33 to Equation 5-36 yield the same form as the constraints for the corresponding wavelength unconstrained Max-RWA ILP model, as expressed in Equation 5-28 to Equation 5-31. Hence, the solution from wavelength constrained source-based model, $S_{constraints}$, is also a solution of wavelength unconstrained source-based model, proving Equation 5-32. What's more, the same conclusion can be deduced similarly for path-based and flow based formulations.

Hence, the optimal objective value derived from the first stage wavelength unconstrained ILP model provides an upper-bound for the final wavelength constrained max-RWA problem.

5.3.3.2 The Second-Stage Graph Colouring Technology

The second stage uses the resulting route-set from the first stage to construct an un-directed colouring graph, and solves the wavelength assignment problem via the graph colouring technology.

The principle of graph colouring is illustrated in Fig 5-5. More specifically, the WDM network is converted to a colouring graph according to the rule that each route in the WDM network is presented by a vertex in the colouring graph, and secondly, any two routes with joint links are neighboured in the colouring graph. When the colouring graph is coloured, neighboured vertices cannot be assigned the same colour.

After the completion of graph colouring, the vertices in coloured graph corresponds to the lightpaths being setup and the colour of each vertex in the coloured graph refers to the wavelength assigned to the corresponding lightpath.

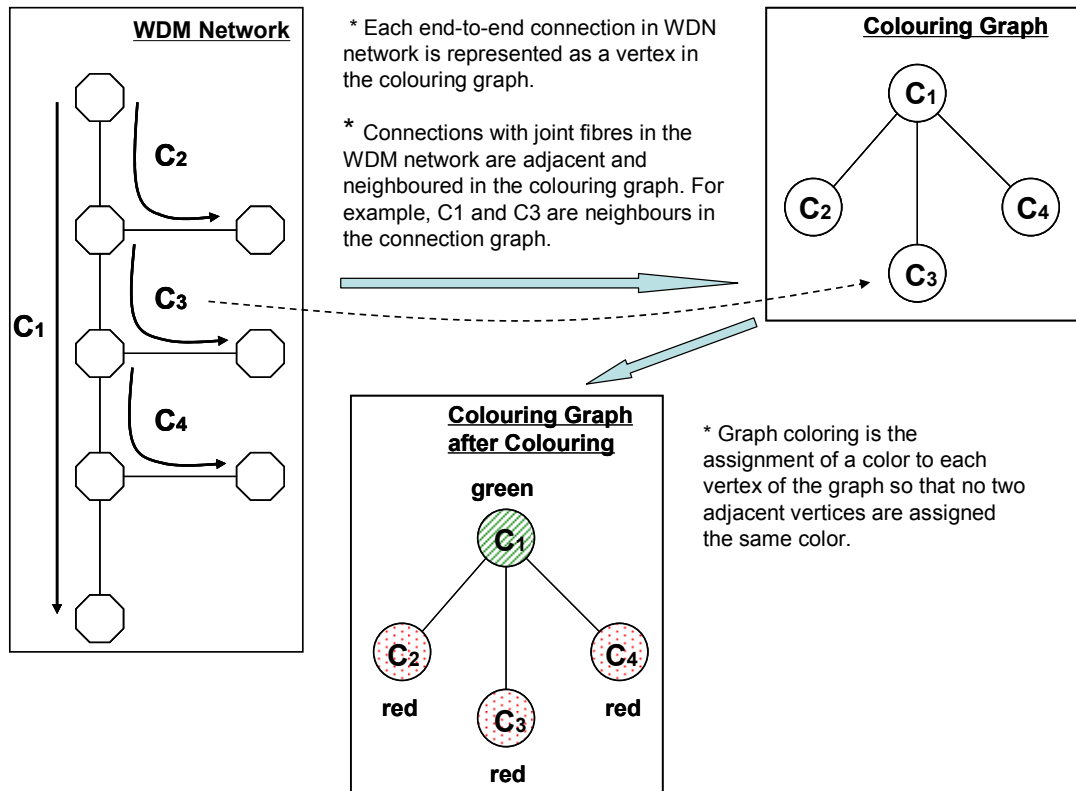


Fig 5-5: Demonstration of Graph Colouring

Normally, the graph colouring algorithms aim to colour the vertices with the least number of colours, and the problem is proved to be NP-complete [BJH02]. In this research, the objective is to colour as many vertices (lightpaths) as possible with the given number of colours (wavelengths). Therefore, if the given number of colours is large enough, the algorithm can stop when all the candidate routes are successfully coloured, whilst, if the number of available colours is small, the algorithm requires more effort to achieve more established lightpaths.

5.3.3.2.1 Sequential Graph Colouring

Graph colouring can be conducted sequentially. The algorithm takes each vertex in a predefined order and tries to colour the vertex with one of the colours used so far. If it is not possible to colour the vertex with any existing colour, a new colour class is created and assigned to the vertex.

There are various sequential graph colouring algorithms, which mainly differ in the vertex colouring order. For example, the vertices can be coloured by an order of descending degree. The order can also be dynamically updated after each step of colouring [Daniel79].

A repetitive colouring mechanism has been proposed to improve the colouring [CL96]. In this mechanism, after the first round of colouring, the coloured vertices are grouped based on the assigned colours. These vertices are sequentially coloured again group-by-group. Such an adjustment in the colouring order can result in a new colouring that requires equal or less number of colours than the previous round of colouring. The proof is very straightforward. Clearly, the vertices in current colour class i can be re-coloured with the same colour i , so the new round of colouring will at most use the same number of colours as the previous round of colouring. What's more, it is possible that some of the vertices in the colour class j can be moved to other existing colours, as long as there is no neighbour relationship between the vertices to be moved and the vertices in the targeted colour classes. In case that all the vertices in one of the current colour classes are relocated to other colour classes, the number of consumed colours is reduced.

The repetitive colouring mechanism rearranges the colouring iteratively, and it guarantees that the resulting number of colours required by the new round of colouring will be no more than that of the previous round of colouring. The difficulty is that the speed of colour number reduction is not predictable, and it may take a significant number of rounds of repetition where the number of colours does not decrease.

5.3.3.2.2 Graph Colouring with Backtracking Enabled Search

Using sequential graph colouring algorithm, the number of required colours are largely determined by the order in which the vertices are coloured. However, the sequential algorithm can work with a proper Depth-First Search (DFS) algorithm, where backtracking and forwarding is employed to find a valid colouring arrangement with the given number of colours³⁶.

³⁶ Instead of finding the minimum number of colours, the objective of the second stage is to find a colouring solution where all the candidate routes derived from the first stage is coloured with the given number of wavelengths (colours). The algorithm ceases when such a solution is found.

More specifically, being applied in graph colouring with k available colours, the vertices are firstly sequentially coloured according to a pre-determined order. When the vertex requires the $k+1^{th}$ colour, the backtracking takes place. The algorithm traces back to the previously coloured node to see whether it can be assigned with an alternative colour. The backtracking continues until one of the previous coloured nodes can be coloured by a different colour, and the colouring proceeds forwardly again to re-colour the next node. Such a forwarding and backtracking procedure happens recursively, until all the vertices are coloured, or until no alternatives can be attempted and k colours are therefore not sufficient for all the vertices.

The above backtracking procedure, based on successive parent vertices tracing, are called chronological backtracking [PGP93]. Fig 5-6 provides an example of DFS search with chronological backtracking, where the vertex ordering is (vertex A, vertex B, vertex C), and two colours, blue and red, are available colours. The numbered arrows demonstrate the progress of the procedure, and the procedure is assumed to get impeded at vertex C after step 4 and step 8³⁷, and backtracking takes place.

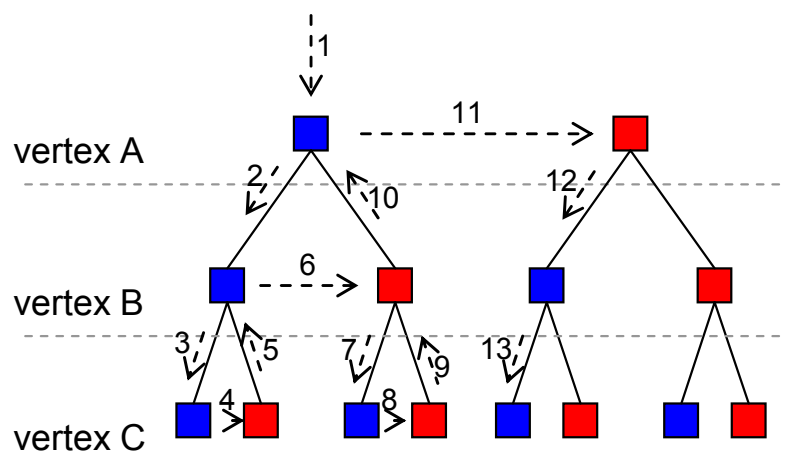


Fig 5-6: An example of BB search with chronological backtracking

5.3.3.2.3 Improvement to Chronological Backtracking

This research proposes an improvement to chronological backtracking for graph colouring, where the ancestor vertices are “jumpily” traced instead of being successively backtracked.

³⁷ This is not a reasonable assumption. However, for simplicity purposes, this assumption is adequate to show the DFS search with chronological backtracking method.

The improvement is based on the following two rules. To assist the rule explanation, Fig 5-7 shows an example graph, where eight vertices (1, 2, ..., 8) are required to be coloured, four colours (A,B, C, D) are available for colouring, and the order in which the vertex is coloured is fixed at (vertex1, vertex 2, ..., vertex 8). Moreover, Fig 5-7 represents an initial colouring, where each node is associated with its vertex name³⁸ and the initially assigned colour (indicated in bracket), and vertex 8 is the problem vertex that can not be coloured by any available colours.

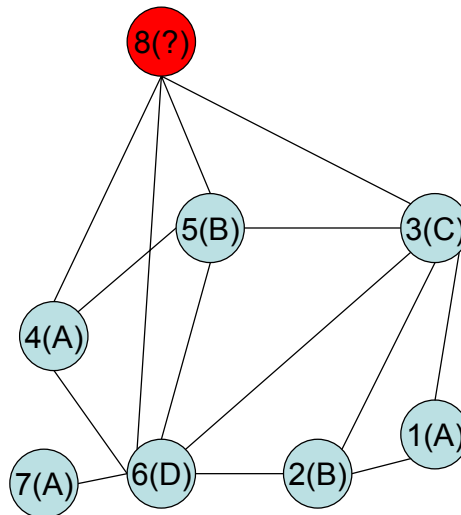


Fig 5-7: An example of colouring graph

Rule 1: When the forwarding colouring is suspended at a certain vertex and the backtracking needs to take place, the problem vertex is actually only affected by its previously coloured neighbours. In other words, re-colouring any un-neighboured ancestor vertices without changing the colour arrangement for neighbored ancestor vertices will not help the suspended vertex to obtain an applicable colour. For example, in Fig 5-7, successively tracing back to vertex 7 and re-colouring it does not help vertex 8 to obtain a valid colour. Therefore, the algorithm can directly trace back to one of the neighbored ancestor vertices to see if the re-colouring is applicable. Such a “jumpily” backtracking strategy saves time from unnecessary repetition of backtracking and forwarding colouring.

Rule 2: A vertex becomes the problem vertex because of the absence of an applicable colour, which further indicates that its neighbored ancestor vertices must have consumed all the

³⁸ The vertex name also indicates the order in which the vertices are sequentially coloured.

available colours, k . Therefore, the problem vertex can only be successfully coloured when the total number of colours used by its neighboured ancestor vertices is reduced to a number less than k . Thus, the backtracking and forwarding colouring should avoid the permutation of colours among those neighboured ancestor vertices. Specially, when these neighboured ancestor vertices form several fully connected cliques with some joint vertices connecting each other, the backtracking point should not start from any of the joint vertices. This is because the re-colouring any of the joint vertices can easily lead to an internal colour swap between those neighboured ancestor vertices, being not able to reduce the number of colours used. The avoidance of choosing the joint vertices as the start of the backtracking point prevents permutation of colours among the neighboured ancestor vertices, thus saving times. Taking Fig 5-7 as an example, directly tracing backing to vertex 6 or 5 will cause colour swapping between vertex 6 and 5, whilst, choosing vertex 4 as a start point of backtracking is an economical way to find an applicable colour for vertex 8.

The above two rules forms an intelligent backtracking for graph colouring. The algorithm examines the ancestor neighbours by a descending order of their colouring tree index. If the neighbour does not belong to any of the joint vertices of multiple cliques formed by those neighbours, the algorithm jumps back to this neighbour directly and continues the forward colouring. However, if the neighbour is a joint vertex between the cliques formed by those neighbours, the algorithm will trace back to the next ancestor neighbour³⁹, and repeat the same backtracking routine until a start point of backtracking is found.

5.3.3.2.4 The Application of Graph Colouring

In the proposed two-stage RWA infrastructure, a sequential colouring is firstly carried out to obtain an initial colouring, and the number of colours, denoted as i , required to setup all the candidate routes derived from the first stage. If i is larger than the provided number of colours, λ , the backtracking graph colouring is employed with the target number of colours setting to $i-1$. The improved backtracking graph colouring is executed iteratively with a decreasing targeted

³⁹ Ignoring some un-neighboured ancestor vertices in this case is effective, because the joint ancestor neighbour between cliques can only change its colour by changing other ancestor neighbours colour, being irrelevant to un-neighboured ancestor vertices.

number of colours, until the entire candidate routes are coloured with λ colours, which means the optimum solution is achieved. Alternatively, the backtracking graph colouring procedure stops in case that the given number of colours is not sufficient to colour all the vertices. In the later case, only part of the candidate routes can be established, and the routes in the largest λ colour classes are selected as the finally supported lightpaths. The number of established lightpaths can be compared to the upper-bound to determine the distance between the current solution and the upper-bound.

5.4 Simulation Models for Proposed Network Architecture

The OPNET™ based simulation models mainly consist of a central network planning node, the core optical switches, and the edge entities.

As shown in Fig 5-8, the “central” node stands for the central network planning node, the numbered nodes represent the core optical switches, and the numbered nodes with a prefix “s” are the edge entities.

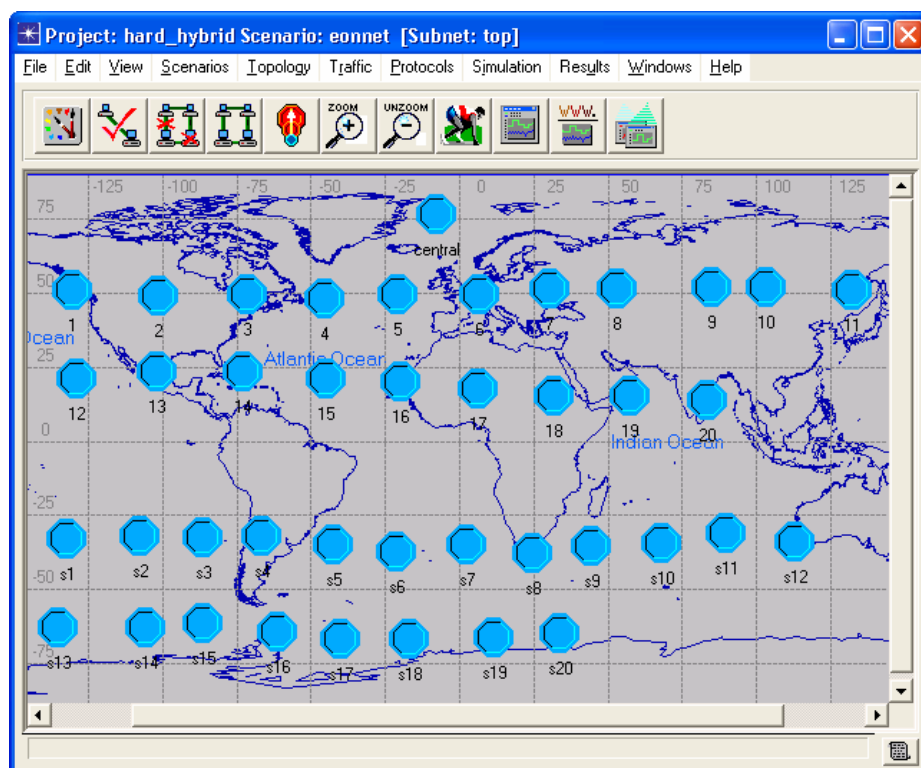


Fig 5-8: Network Structure of Proposed Hybrid Network Architecture

5.4.1 Central Network Planning Node

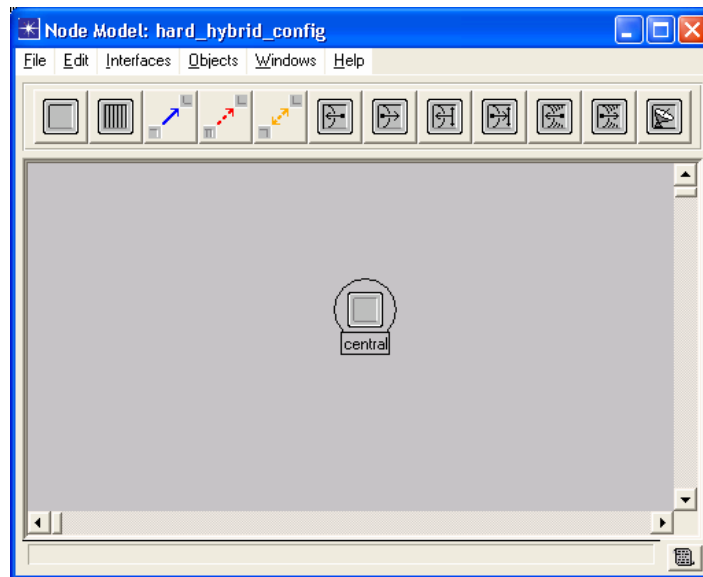


Fig 5-9: Node Model of the Central Network Planning Node

The central network planning node corresponds to the central node owned by the CNO. As shown in Fig 5-9, the structure of the central node is rather simple as a single processor, which is able to communicate with all the other nodes.

The main functions of the central node are to accept lightpath pre-booking requests from various SPs, and run an RWA algorithm to decide the lightpath placement for those predicted lightpath demands. After the lightpath placement solution is produced, the central node sends signals to the corresponding edge and core nodes to setup the wavelength-routed end-to-end lightpaths.

In terms of the RWA algorithm employed in the simulation, the central node uses a weight-updated Dijkstra algorithm to calculate route and wavelength for each pre-booking request. More specifically, the weight-updated Dijkstra algorithm is one of the sequential RWA approaches, where the lightpaths are calculated in sequence. The algorithm uses an undirected weighted graph $G=\{V, E, W\}$ to represent the network topology, where V is the set of network vertices, E is the set of unidirectional links, and W is the set of weights on links. For each lightpath to be setup, the algorithm firstly runs the Dijkstra algorithm to calculate a route with the least weight, it then goes through the links along the route to find a wavelength for the lightpath. When both the route and wavelength are successfully found, the link weights, and wavelength deployment information along the path are updated, and the algorithm is executed

again for the next request. If all the wavelengths on the link are consumed, the link weight will be set to infinity, such that the link is dropped from the graph. In the simulation, the link-distance is used to reflect the link weight, where the longer is the link the larger will be the link weight value.

As it is hard to find a network-wide traffic matrix that varies with time, each round of simulation is assumed to be a pre-booking cycle, where the pre-booking requests are collected at the beginning of the simulation, and the pre-booking demands are not varied during the simulation period⁴⁰.

5.4.2 Edge Nodes

As the edge and core optical switch nodes deal with both wavelength routing and conventional OBS, they were built based on the models for the classical OBS networks, as shown in Chapter 4, section 4.2.2.

The edge nodes generate burst traffic for all the possible destinations, send pre-booking requests to the central node, receive pre-booking acknowledgements from the central node, and decide whether the bursts should be delivered on pre-booked end-to-end lightpaths or using a classical OBS service. In addition, the edge node is also responsible for calculating a route from itself to various destinations, such that a source-routed routing table is maintained at the edge SP node for OBS routing. In implementation, a shortest path is calculated and stored in the OBS routing table.

More specifically, apart from the functionality blocks for the classical OBS operations, the edge nodes maintain a pre-booked lightpath record, which keeps all the pre-booked resource information. When a new burst arrives, the edge node looks up those pre-booked resources, and decides whether the burst can be sent on one of the pre-booked lightpaths. If there are no appropriate reserved resources for the burst, the edge node will issue a typical OBS operation to ship the burst with a risk of being lost on the way.

Furthermore, when the edge node decides the allocation of pre-booked resources for bursts, it

⁴⁰ In reality, the central node actively reacts to the updated pre-booking demands, and rearranges the lightpath placement accordingly.

employs the Latest Available Unused Channel with Void Filling (LAUC-VF) algorithm to select the wavelength path, such that the reserved resources are efficiently utilized [XVC00]. As to the wavelength selection for OBS services, the LAUC-VF algorithm is adopted in case that the wavelength conversion is allowed for the OBS operation, whilst, the wavelength is randomly selected at the edge node in case that the wavelength conversion is not allowed for the OBS service. This is because if the wavelength conversion is not supported for the OBS reservations, the wavelength is unchangeable at any intermediate nodes. Even if the optimal wavelength is selected on the first link using the LAUC-VF algorithm, it cannot be guaranteed that the selected wavelength is the best one on later links.

5.4.3 Core Optical Switch

The core optical switches react to the signals sent from the central node, and setup pre-booked lightpaths. The core switch nodes also relay the bursts sent on the pre-booked lightpaths, or deal with the typical OBS operations.

In a core optical switch node, a wavelength routing table is established and maintained. The wavelength routing table mainly records the incoming and the outgoing ports for the pre-booked lightpaths, and the bursts being sent on those lightpaths will be transparently switched from the incoming port to the outgoing port. The classical OBS operation in terms of the BHP manipulation and the burst data relaying are the same as the old JET model.

5.4.4 Validation of the Simulation Models

The models are validated at each key stage of the modelling.

Firstly, the network model with only the classical OBS is checked. The activities of the BHP and the burst data are closely observed, and it is confirmed that the classical OBS functionalities operate correctly in the model. Secondly, the central node is added to the simulation. As the central node mainly contains some function blocks, each function is debugged offline to ensure all the functions perform correctly. Thirdly, the establishment and the maintenance of the pre-booking record in the edge node, and the wavelength routing table in the core optical node are verified. The methodology used is the same as the one used in the OBS network, which prints out every change and status of the tables, and to ensure the corresponding records and

tables are maintained and updated accurately. Finally, the whole network with both the wavelength routing and the OBS are checked by observing the burst activities between nodes.

5.5 Simulation Results for Proposed Network Architecture

In simulations, the USA NSFnet is used as the network topology, as shown in appendix B. 32 wavelengths are assumed on each unidirectional link, and each wavelength is operating at 10Gbps. When the bursts are generated, they are emitted at a constant interval, in which the burst aggregation time is set to 25 milliseconds. According to ref [YLC04], the burst length distribution using the time-based aggregation mechanism is close to the Gaussian distribution, so the burst length is a randomly generated number according a normal distribution. The same amount of traffic is generated for all the source-destination pairs, and the traffic load used in the resulting figures refers to the ratio of average incoming bit rate to the wavelength capacity for each source-destination pair.

When the SP sends the pre-booking request, it adopts the lightpath reservation strategy as shown in Fig 5-3. Therefore, when the predicted load for a source-destination pair is within the range of 0.0 (exclusive) to 1.0 (inclusive), one lightpath will be pre-booked for the specified pair. Similarly, when the predicted load ranges from 1.0 (exclusive) to 2.0 (inclusive), two lightpaths will be reserved. In simulations, the traffic prediction is assumed to accurately reflect the actual incoming rate for each source-destination pair.

To make the simulation more realistic, each established lightpath for pre-booking requests has to yield a unique wavelength throughout, such that no wavelength conversion is required. However, for classical OBS operation, both the wavelength conversion supported and un-supported scenarios were carried out. The performance is mainly illustrated by the blocking probability with varying load and maximum allowed delay parameters.

5.5.1 Blocking Probability versus the Load

In this set of experiments, the traffic load varies from 0.1 to 2.0, where the number of pre-booked lightpaths for each source-destination pair is 1 when the load is no more than 1.0, and 2 when the load is between 1.0 and 2.0.

Burst data can be further delayed at the edge in case the pre-booked resources are not immediately available. The maximum allowed delayed time is denoted by the “max_delay” as shown in the resulting figures.

For comparison purposes, the performance of a pure classical OBS network and the pure static Wavelength Routed (WR) network are presented together with the proposed network architecture. The result for pure classical OBS network does not allow any further edge delay⁴¹, whilst, that for static WR networks allows some extra edge delay as in the proposed network architecture.

5.5.1.1 Results without Wavelength Conversion Support for OBS Services

As wavelength conversion is expensive, this scenario investigates the blocking probability of the proposed network architecture when there is no wavelength conversion for either the pre-booked lightpaths or the OBS service.

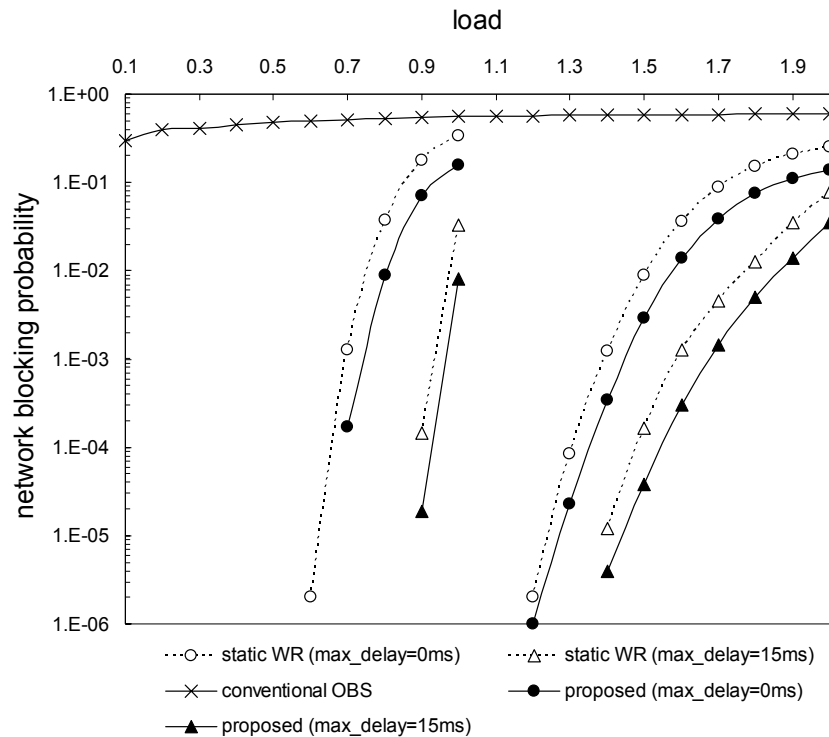


Fig 5-10: Blocking Probability versus the Load in a Network without Wavelength Conversion Support

⁴¹ This is because the classical OBS does not provide any admission control functionality.

The network burst blocking probability is illustrated in Fig 5-10. The result shows that for a network with pure OBS operations without wavelength conversion support, the blocking probability is extremely high. The classical OBS technology usually requires the wavelength conversion support at every network node, such as to increase the bandwidth and achieve acceptable blocking probability. The classical OBS without wavelength convention suffers from un-acceptably high loss rate.

The proposed hybrid network architecture yields lower blocking than the pure OBS network. When the load is under 1.0 and one lightpath is pre-booked for each source-destination pair, nearly up to 0.7 traffic load can be supported with a blocking rate less than 10^{-4} without experiencing any extra edge delay ($\text{max_delay} = 0$ ms). Similarly, a traffic load of nearly 1.35 can be supported with a blocking rate less than 10^{-4} , when the load is between 1.0 and 2.0, and two lightpaths are pre-booked for each source-destination pair. This is because the proposed network provides two approaches to deliver burst traffic. The pre-booked lightpaths sends the traffic with guaranteed end-to-end reachability, and the classical OBS provides best-effort services if the burst cannot be fed into the pre-booked resources.

The performance of static WR provides a means to show the contribution made by the OBS services in the proposed network architecture. It can be seen that the OBS services start to show benefits when the traffic load gets to the higher end of the pre-booking regime. This is because that most of the traffic can be successfully transmitted on the pre-booked lightpaths, when the traffic load is low, thus the OBS services are rarely utilised. When more traffic comes into the network, the pre-booked resources are not sufficient, so the OBS services begin to help and deliver burst data with its inherent risk of losing data on the way.

Furthermore, the results also show that the extra edge delay (max_delay) can improve the blocking probability performance. With the proposed network architecture, the supported traffic load increased up to nearly 0.95 when 15 milliseconds extra edge delay is allowed for the bursts with one lightpath pre-booked for each source-destination pair. Similarly, the supported traffic load goes up to nearly 1.55 when two lightpaths are pre-booked for each source-destination pair. The extra edge delay allows the burst to wait at the edge node until the pre-booked resources are available at a later time, which also helps to fill the pre-booked resources with more burst

traffic.

5.5.1.2 Results with Wavelength Conversion Support for OBS Services

This experiment assumes that each pre-booked lightpath still yields a unique wavelength throughout, but the wavelength conversion is allowed for classical OBS services at any network node. Other experimental parameters are same as those in section 5.5.1.1. The result is presented in Fig 5-11.

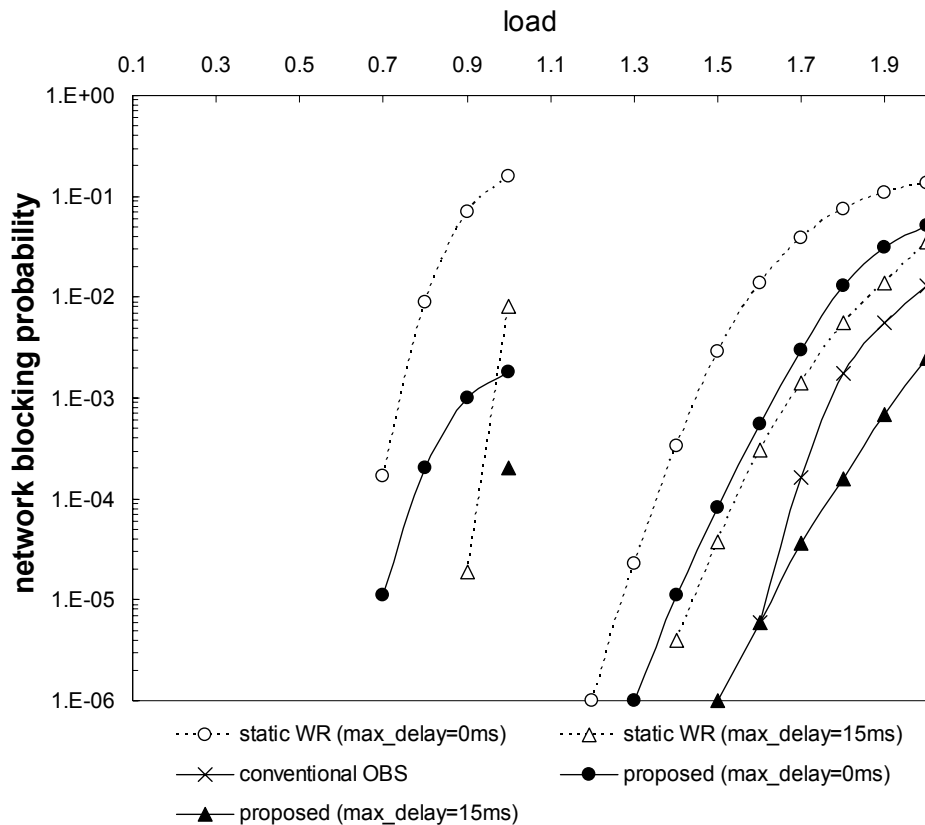


Fig 5-11: Blocking Probability versus the Load in a Network with Wavelength Conversion Support for OBS Service

Unlike the case without wavelength conversion support, the pure classical OBS with wavelength conversion support achieves a very low blocking probability with 32 wavelengths provided on each link. It yields non-blocking performance until the traffic load is higher than 1.5. Therefore, conventional OBS technology is feasible when the wavelength conversion technology matures and becomes readily available.

However, the proposed hybrid network architecture and the corresponding static WR suffers more

blockings than conventional OBS, unless the proposed architecture allows 15 milliseconds extra edge delay at a load higher than 1.5. This indicates that, in NSFnet, with 32 wavelengths on each link and wavelength conversion support for OBS services, the proposed network architecture does not perform better than classical OBS. This is because the wavelength conversion can increase the capacity of the classical OBS networks.

5.5.2 Blocking Probability versus the Allowed Extra Edge Delay

Fig 5-12 illustrates the blocking probability for the proposed network architecture, with various values of extra edge delay allowed for burst buffering at the edge. 32 wavelengths are provided on each link, and traffic load is fixed at 1.0.

As expected, the blocking probability decreases when more delay is permitted. Additionally, if the network provides wavelength conversion for the OBS operation, the blocking probability will be lower than that without wavelength conversion support.

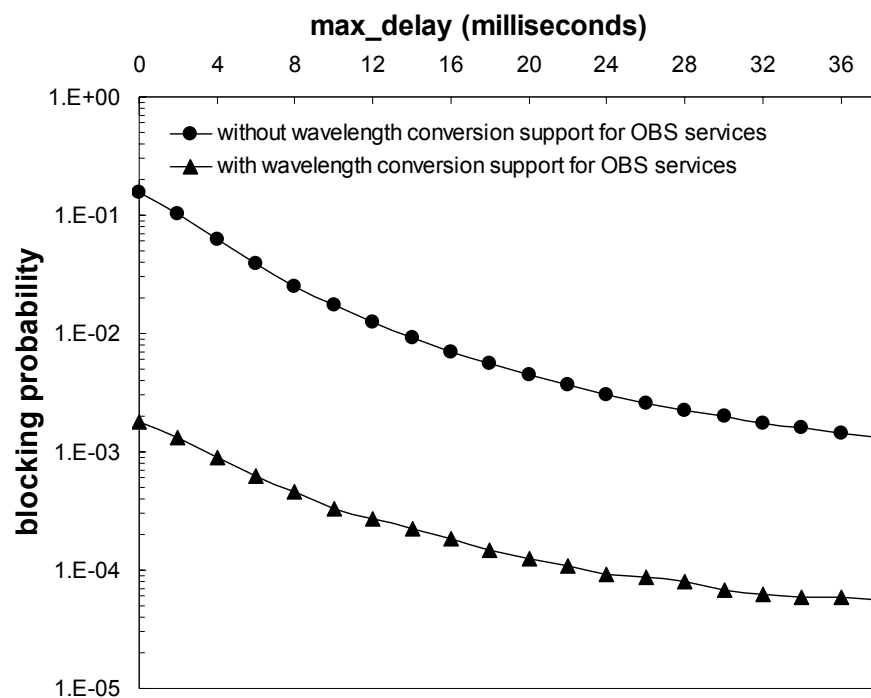


Fig 5-12: Blocking Probability versus the Allowed Extra Edge Delay

5.6 Numerical Results for ILP-based RWA Models

This section firstly compares and analyses the numerical results derived from the conventional

wavelength unconstrained ILP models, followed by the results for proposed two-stage RWA approach. The results for conventional wavelength constrained ILP models are not presented, because those models cannot provide any solution after running for several days.

In terms of implementation, ILP formulations are modelled and solved by the “lp-solve” software. The experiments are run in two networks, the 14-node USA NSFnet and the 20-node European EONnet, respectively. Additionally, the frequently cited traffic matrix as used in references [KS01][BCB04][JMT04] is employed in the experiments, in which 268 lightpath demands ask to be established in NSFnet, and 374 lightpath demands are required to be setup in EONnet. The details of these matrices are presented in appendix C.

5.6.1 Results of Conventional ILP Models without the Wavelength Continuity Constraint

Table 5-4 illustrates the objective values derived from the corresponding wavelength unconstrained ILP models in NSFnet and EONnet with the traffic matrices as in appendix C. As the path-based ILP model requires all the loop-free routes as the input, a total of 13187 routes are calculated for NSFnet. For EONnet, the total number of possible routes are much higher, so the routes provided in the model is restricted to those with at most 6 hops, and a total of 12014 routes are offered in the test.

Table 5-4: Objective Value from the Wavelength unconstrained ILP Models

Number of Wavelengths	Objective Value					
	NSFnet			EONnet		
	path	flow	source	path	flow	source
10	198	198	198	285	285	285
11	208	208	208	301	301	301
12	218	218	218	317	317	317
13	228	228	228	329	329	329
14	238	238	238	337	337	337
15	248	248	248	344	344	344
16	258	258	258	350	350	350

17	263	263	263	356	356	356
18	267	267	267	362	362	362
19	268	268	268	367	367	367
20	268	268	268	370	370	370
21	268	268	268	373	373	373
22	268	268	268	374	374	374

The table clearly shows that the objective values from the three models are the same for both the NSFnet and the EONnet⁴². As there are totally 268 lightpath demands for NSFnet and 374 lightpath demands for EONnet, all the three ILP models are able to achieve non-blocking probability performance when at least 19 wavelengths are available in NSFnet and 22 wavelengths in EONnet, subject to wavelength conversion being supported on every node of the network. In addition, the results also indicate that in EONnet, only providing part of the possible routes with maximum 6 hops is sufficient for producing the optimal solution with path-based formulations. However, it is hard to prove the optimality without comparing the result to the other models.

Table 5-5 demonstrates the time taken to produce the solution for the above experiments on a 3.06GHz Pentium 4 personal computer. It shows that in NSFnet, path-based and flow-based models yield comparable performance in terms of the solving speed. In EONnet, the flow-based model is faster than the path-based model. Most importantly, the source-based model achieves the shortest problem solving time among the three models in both networks. For all the cases being considered, the source-based model solves the problem within milliseconds to a few seconds.

Table 5-5: Execution Time for the Wavelength unconstrained ILP Models

Number of Wavelengths	Execution Time (in Seconds)	
	NSFnet	EONnet

⁴² The details of the exact solution from the three models may differ from each other if multiple solutions are available.

	path	flow	source	path	flow	source
10	6.765	6.344	0.062	221.833	83.481	0.453
11	6.578	5.843	0.172	117.897	86.543	1.406
12	4.624	6.656	0.125	157.926	73.918	0.39
13	3.327	5.64	0.093	132.632	60.716	0.36
14	9.374	3.703	0.078	236.829	58.279	0.578
15	10.094	4.015	0.078	1.453	44.951	0.657
16	3.922	5.015	0.078	139.523	40.905	0.219
17	50.467	3.343	0.079	252.626	37.233	0.187
18	0.891	3.015	0.063	1.656	25.655	0.219
19	0.765	2.656	0.078	212.409	27.436	0.297
20	0.704	2.187	0.047	235.565	22.592	0.203
21	0.688	1.875	0.062	1.078	21.046	0.172
22	0.687	1.531	0.063	273.703	18.359	0.172

5.6.2 Performance for the Proposed Two-Stage Approach

The direct wavelength constrained ILP model requires large amount of computational resources to produce the optimal solution. When the problem size gets large, the direct ILP approach easily becomes impractical. The proposed two-stage approach for wavelength constrained Max-RWA problem suggests that the problem can be firstly solved in a more general manner, where the wavelength conversion is temporarily allowed. After candidate route generation takes place from one of the wavelength unconstrained ILP models, graph colouring technology is employed to impose wavelength continuity constraint and assign a unique wavelength for each established lightpath. It has been proved that the objective value from the first stage ILP model is the upper-bound of the final solution. Thus, the final solution can be compared to the upper-bound to get an idea of optimality performance⁴³.

In terms of implementation, the source-based wavelength unconstrained ILP model is employed

⁴³ There has not been any practical methodology available to calculate the exact optimal value for the wavelength constraint Max-RWA problem. Therefore, the degree of optimality can only be presented by the upper-bound value.

in the first stage. The improved backtracking graph colouring algorithm is used in the second stage, and the results for the specified traffic matrices in 14-node NSFNET and 20-node EONNET are illustrated in Table 5-6.

Table 5-6: Performance of Proposed RWA Mechanism

wavelengths	NSFNET				EONNET			
	B1	B2	C	T	B1	B2	C	T
10	198	198	198	0.094	285	285	285	0.625
12	218	218	218	0.172	317	317	317	0.672
14	238	238	238	0.203	337	337	337	1.5
16	258	258	258	0.109	350	350	350	0.422
18	267	267	267	0.109	362	362	362	0.39
20	268	268	268	0.094	370	370	370	0.422
22	268	268	268	0.078	374	374	374	0.406
24	268	268	268	0.094	374	374	374	0.410

Notations:

B1: the upper-bound value derived from the direct source-based wavelength constrained Max-RWA ILP model, where LP-relaxation is applied.

B2: the objective value derived from the first stage ILP model in the proposed approach.

C: the number of finally established lightpaths using the proposed approach.

T: the time (in seconds) used to derive the solution in the proposed approach.

The results clearly show that in all the cases being studied, the upper-bound derived from the source-based wavelength unconstrained Max-RWA ILP is the same as that from the corresponding wavelength constrained ILP model with LP relaxation, that is, B1=B2.

Encouragingly, with the proposed two-stage RWA mechanism, the upper-bounds for the given matrices are reachable, which means that the optimal solution can be found for the matrices under study.

More importantly, the proposed mechanism exhibits ultra high solving speed. For the first time, the problems are solved with confirmed optimality within milliseconds to a few seconds. However, the best results published to-date for the same matrix instances are at a timescale of

minutes as in ref [BCB04][JMT04], where the ILP is solved by a more advanced commercial solver CPLEX-MIP.

The proposed mechanism can guarantee that the obtained solution is optimal when the second stage is able to colour all the candidate routes derived from the first stage. However, if the second stage fails to colour all the candidate routes, there is no proof of optimality for the obtained solution, because the exact optimal objective value cannot be revealed with current technologies. Nevertheless, the upper-bound value will be able to provide an idea of how good the solution is from the upper-bound.

5.7 Summary

This chapter proposes and investigates a new hybrid network architecture, where the pre-booking services are provided for large timescale average load prediction. The SPs and the CNO cooperate to establish end-to-end wavelength routed lightpaths for predicted traffic loads. Meanwhile, the SPs and the CNO can also arrange a typical classical OBS signalling to deliver the bursts with the best efforts.

The simulation results show that in the USA NSFnet with 32 wavelengths on each link, if no wavelength conversion support is available to either the WR lightpaths or the OBS services, the proposed network architecture achieves lower blocking probability than that of pure classical OBS. The OBS services in the proposed architecture help to reduce the blocking probability, when the traffic load gets high and the reserved WR lightpaths become insufficient for the traffic load. However, if the wavelength conversion is allowed for the OBS services, the proposed architecture does not show any advantages over the pure OBS. This indicates that the classical OBS technology is promising when the wavelength conversion technology becomes mature.

This chapter also studies various ILP formulations for the wavelength constrained Max-RWA problem, which play an important role in achieving the minimum blocking probability for WR networks. The analytical and the experimental studies show that the direct ILP model for the wavelength constrained Max-RWA problem consumes a large amount of computer resources, thus taking long time or even being impractical. This research proposes a two-stage RWA

algorithm, and an improvement to the traditional backtracking graph colouring methodology, such that the search space is dramatically reduced and the problem solving speed is accelerated.

The numerical results utilising the source-based ILP model are very encouraging. For the first time, confirmed optimal solutions for the traffic matrices under study are achieved by the proposed methodology within a short time-frame. The problem solving time for all the cases are within milliseconds or to seconds region on a standard personal computer. Further research is required to verify the generality of the proposed RWA approach for any networks with various traffic matrices.

Chapter 6 Pre-booking Mechanism for End-to-End Connectivity Acknowledged OBS Networks

6.1 Introduction

In end-to-end connectivity acknowledged OBS networks, each burst waits for the lightpath acknowledgement before it is admitted to the core. WR-OBS [DB02][DB022] and MPLS/GMPLS-based OBS [SST03][STS04] are two good instances, and they employ a two-way reservation protocol to establish an end-to-end lightpath for each burst. Once the lightpath is established, the burst can be transparently delivered from the source to the destination, so traffic loss only occurs when insufficient resources are found and no lightpaths can be established during the lightpath setup stage⁴⁴, or when the edge buffer overflows. If the edge buffer size is infinite, the traffic loss is only caused by the absence of network resources⁴⁵. Therefore, it is very important to utilise the network resources in an efficient manner, such that the network can support the maximum volume of traffic.

However, experiments in WR-OBS show that the two-way reservation protocol causes a low lightpath bandwidth efficiency, where the lightpath bandwidth efficiency is defined as the ratio of burst transmission time to the lightpath holding time, and consequently impedes the end-to-end delay guaranteed service provisioning in large backbones [MKB02].

In fact, in traditional two-way reservation protocols, the lightpath is usually configured and set to be held during the lightpath acknowledgement phase, and the release of the lightpath has to wait until the corresponding lightpath teardown message is received. As the signalling requires time to propagate from node to node, some idle resources have to be held longer than necessary. If the burst transmission time is short and the signalling time is long, only a small portion of reserved resources is consumed by the actual burst transmission, whilst the other part is reserved

⁴⁴ The traffic loss due to physical layer faults, such as the fibre cuts, the network node failure, etc, is not considered here.

⁴⁵ This study assumes the edge buffer size is infinite, and concentrates on the efficiency of network resources and its effects on traffic loss performance.

for idle holding, so this type of reservation will yield low lightpath bandwidth efficiency, and degrade the network performance, in terms of blocking probability and so forth.

To this end, the concept of “delayed reservation” (DR) in JET-based OBS [CQY04] appears very attractive in terms of the efficient resource reservation. DR makes use of intelligent software scheduling to control the OXC, and the OXC is only configured during the burst transmission time, excluding any unnecessary resource consumption.

This study inherits the spirit of DR and proposes a novel pre-booking mechanism to provide an efficient network architecture for future WDM networks, where the end-to-end connectivity acknowledged reservation protocol is employed for lightpath provisioning, but the lightpath bandwidth efficiency is improved by the pre-booking mechanism.

In the proposed pre-booking mechanism, the lightpath configuration is proactively scheduled by either explicit booking or implicit burst prediction, excluding any unnecessary resource holding caused by the signalling propagation. Additionally, in order to promptly withdraw the lightpath when the burst transmission completes, burst length prediction is required to estimate the time to release the resource. As self-similar traffic is widely observed in the Internet [EA97][ESW94][ZS99][AMD02], we analyze the aggregated burst length statistics from the self-similar incoming traffic and derive a burst length analysis model for burst length prediction purposes.

6.2 Network Architecture

The proposed network structure is shown in Fig 6-1, where the WDM backbone is operated by a Carrier Network Operator (CNO), with various Service Providers (SPs) placed at the edge.

SPs manage bursts at the network edge. Assuming a lack of optical buffering facilities in the core, each SP aggregates traffic with the same destination and class of service into bursts at the ingress. Additionally, each SP sends and receives the pre-booking request for each burst, and controls the admission of bursts into the core.

The core network management operates as a multi-service commodity market, allowing various SPs to pre-book network resources at specified times and durations in the future. Meanwhile,

via the pre-booking mechanism, the CNO maintains all the committed lightpaths information, and schedules the OXC configuration so that the necessary switching actions can be taken at the required time(s), and the data transport occurs transparently in the CNO optical layer based on the committed reservations.

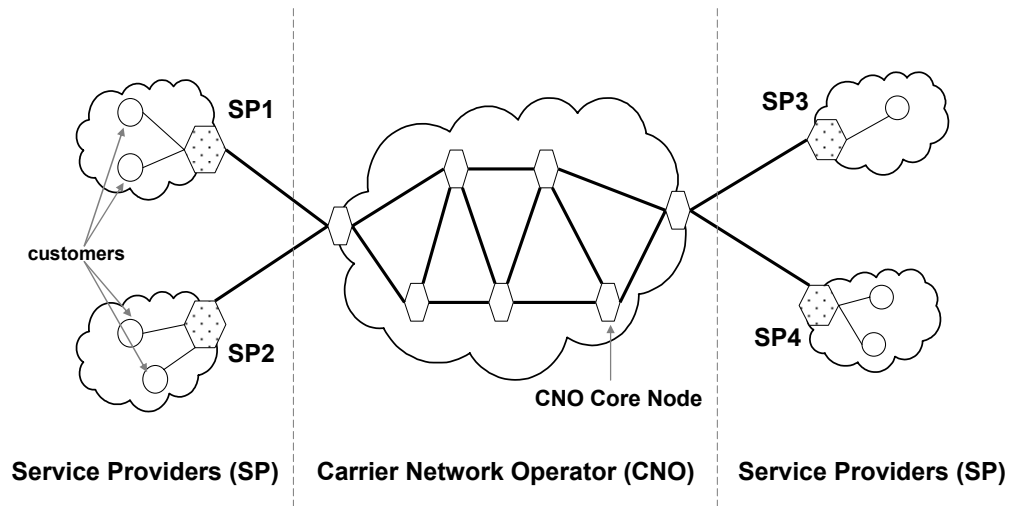


Fig 6-1: Proposed network architecture

The architecture can be managed by either a centralized control plane, such as the WR-OBS, or a distributed control plane, for example, based on MPLS/GMPLS control. In the centralized control plane, all the requests and actions in the core network are handled by a central node, where the geographically central location is usually chosen for the control node. With distributed control, each node in the core acts independently, and their operation is based on the signalling protocol and other associated policies agreed among all the nodes in the core. Nevertheless, the control plane will not change the benefits of the pre-booking mechanism itself, as the resource saving in the proposition is achieved by reducing unnecessary holding of resources during the signalling phase.

In terms of the hardware requirement, the pre-booking mechanism would require memories to store the pre-booking information, and a controller for each OXC to schedule the configuration of the fibre fabric. As each burst/flow duration in the pre-booking system is of a magnitude comparable to the end-to-end propagation delay, which is in millisecond region (this will be described in more details in a later section), the pre-booking mechanism can adopt the latest MEMS-based mirror substrate in the OXC, which yields a switching speed in millisecond

region. Additionally, the burst/flow granularity in the pre-booking mechanism also relieves the requirements on the processing speed of the control unit, and makes the electronic controller practically adoptable in the optical core.

6.3 Explicit and Implicit Pre-booking

Pre-booking at a SP site includes explicit and implicit pre-booking relationships with its customers.

Explicit pre-booking deals with definite booking requests from customers. It usually reserves resources for longer to accommodate the bulked aggregate traffic from many of its customers, and the reservation is committed a long time before the actual transport takes place. For example, a bank may regularly reserve a considerable amount of bandwidth around 4:30 PM every day to securely deliver internal data to its central archival facility. Conversely, implicit pre-booking accommodates other short duration bursts via traffic prediction. The SP's reservation procedure makes use of estimated burst information to decide the reservation, and provides committed resource information before the transmission starts.

As explicit pre-booking can reserve the network resource long before the transmission starts, there is sufficient time for the network planner to derive a good lightpath placement decision. Meanwhile, the pre-booking only reserves the resources for the duration of actual flow transmission, so no extra resource is idly reserved, leading to a full utilization of the reserved lightpath bandwidth. Additionally, as the network resource provisioning is based on the first-come-first-served principle, the earlier the explicit pre-booking is requested, the higher priority and probability the request can be served.

Implicit pre-booking is actually based on the reserve-while-accumulating idea, such that the burst aggregation and the lightpath reservation are handled in parallel, which reduces the end-to-end delay. Furthermore, the implicit pre-booking adopts intelligent scheduling to avoid holding idle resources during the signalling propagation time, improving the lightpath bandwidth efficiency, hence enhancing other associated measures of performance.

The key challenge is to reveal the benefit margin of implicit pre-booking in conjunction with specific prediction approaches. Therefore, this chapter mainly focuses on implicit pre-booking.

6.4 Burst Assembly Mechanisms in Implicit Pre-booking

Implicit pre-booking takes advantages of burst length prediction and reserves a lightpath for each upcoming aggregated burst, exclusive of explicitly booked traffic.

More specifically, the burst assembly can adopt the Limited-Burst-Size (LBS) or Unlimited-Burst-Size (UBS) mechanism to finalise each burst aggregation [DB02][MDB01]. As the burst assembly should at least last for the signalling round-trip-time, the assembly timescale in this pre-booking mechanism will be comparable to the signalling round-trip-time and is within the millisecond region.

6.5 Implicit Pre-booking Signalling with Centralized Control

With centralised control, the link-state database and the lightpath calculation is maintained and operated at a central control node. As the central node possesses global knowledge of the network link-state, the routing and the lightpath selection will be more flexible and efficient than with a distributed control plane.

6.5.1 Traditional Signalling and Reservation

A good example of the two-way connectivity acknowledged reservation with centralized control is the WR-OBS [DB02].

The burst aggregation starts when the first bit of the new burst arrives. Depending on the burst assembly mechanism (LBS or UBS), the burst aggregation stops when the acknowledgement arrives or when the buffer is emptied. As shown in Fig 6-2, the lightpath reservation starts after the burst is aggregated for an initial time T_{edge} , with a Lightpath Request from the ingress node to the central control node.

In the central node, the requests from various edge nodes are buffered, scheduled and processed using a Routing and Wavelength Assignment (RWA) algorithm with the Earliest Deadline First (EDF) principle. The RWA algorithm calculates a lightpath for the request based on the link-state database information. If no lightpath is available at the time of calculation, the request may be temporarily buffered until some resource becomes available within the delay tolerance constraint. The request will be rejected if no lightpath can be found before the delay tolerance is

violated.

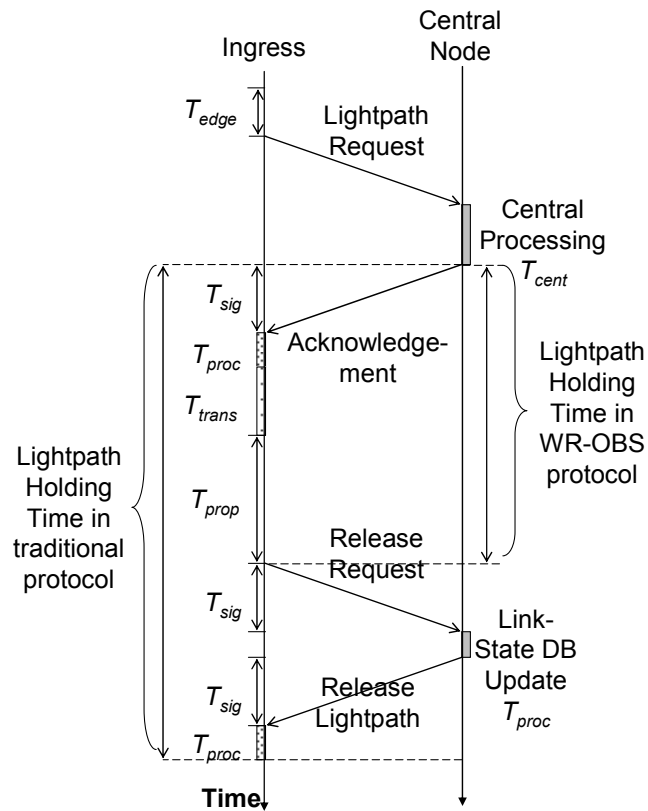


Fig 6-2: Traditional lightpath signalling in centralized control plane

Assuming that a lightpath is found for the request, the involved resources are regarded as being held by the central node, and the central node will then send an acknowledgement to the ingress node confirming the lightpath, and at the same time the corresponding lightpath nodes are configured by receiving the signalling from the central node.

On receiving the acknowledgement, the burst is emitted to the core and is transparently delivered to the other end. All previous WR-OBS research has been based on the perfect knowledge of the burst length, thus the lightpath will be released once the burst reaches the destination. However, this is an over-simplification. In practice, a Release Request should have been sent from the egress node to the central node informing the completion of the delivery, and the centralized node then updates its link-state database and sends Release Lightpath message to each of the lightpath node to release the OXC configuration.

Thus, the lightpath holding time starts when the lightpath is found by the central node, and it is not released until the Release Lightpath message is received by each of the lightpath node and

the OXC is re-configured.

Assuming the burst transmission time is T_{trans} , the burst propagation time is T_{prop} , the OXC configuration time (either setup or release) and the central node database update time are all T_{proc} , and the Signalling propagation time from the OXC to the central node is T_{sig} , then, the lightpath holding time in traditional reservation with centralized control is expressed as shown in Equation 6-1.

$$T_{lightpath}^{tradition-C} = T_{trans} + T_{prop} + 3T_{sig} + 3T_{proc} \quad \text{Equation 6-1}$$

As WR-OBS assumes that the lightpath will be released as soon as the burst reaches the destination, the lightpath holding time in WR-OBS can be expressed as in Equation 6-2.

$$T_{lightpath}^{WROBS} = T_{trans} + T_{prop} + T_{sig} + T_{proc} \quad \text{Equation 6-2}$$

6.5.2 Proposed Signalling and Reservation

With pre-booking, as shown in Fig 6-3, the lightpath pre-booking request is initialized at the very beginning of the burst aggregation. Based on observed traffic characteristics and associated QoS criteria, the request includes the earliest burst transmission time $T_{earliest}$, the maximum tolerable edge delay max_delay , and the predicted burst length information in the reservation request message. After the signalling propagation time, T_{sig} , the request reaches the central node, and is queued into the request buffer. The request will then be scheduled based on the EDF principles and processed using a RWA algorithm. On the completion of RWA, a pre-booking acknowledgement is sent back to the edge node, and the burst transmission happens at the allotted reservation time.

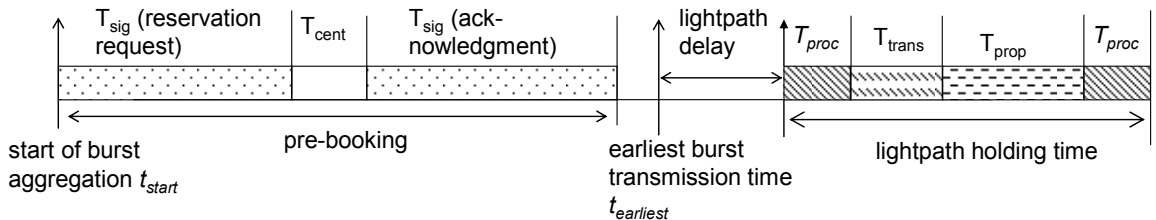


Fig 6-3: Pre-booking signalling time parameters in centralized control plane

The pre-booking has to be completed before the burst transmission starts, so the earliest burst

transmission time should satisfy Equation 6-3. Therefore, the burst aggregation time is at least twice T_{sig} , which will be comparable to the one-way end-to-end propagation time.

$$t_{earliest} \geq t_{start} + 2T_{sig} + T_{cent} \quad \text{Equation 6-3}$$

Fig 6-4 presents a simplified figure on the proposed signalling, which can be used to do comparison with Fig 6-2. As it shows, the lightpath reservation is accomplished using the Lightpath Request and the Acknowledgement messages. However, when the request is processed by the central node, the RWA algorithm is only executed once for each request to find a lightpath possibly with some delay, based on the committed reservation and the delay tolerance provided by the request (more details concerning the RWA algorithm will be described later). The central node keeps all the committed resource information in the link-state database, and sends signalling messages to the corresponding OXCs to perform the setup just before the burst transmission starts at the ingress, and to ensure the resource release after the last bit of the burst reaches the egress node.

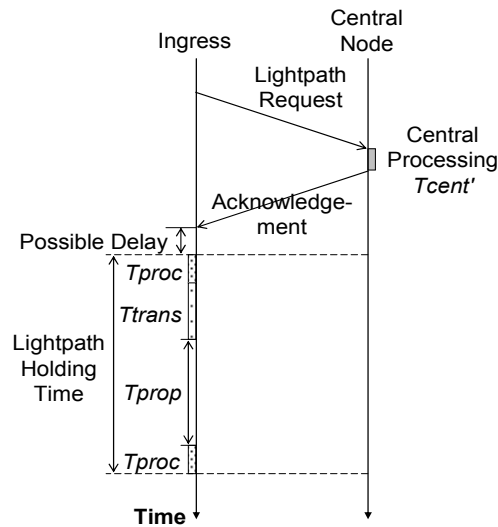


Fig 6-4: Pre-booking Lightpath signalling in centralized control plane

Therefore, the lightpath holding time can be expressed in Equation 6-4.

$$T_{lightpath}^{prebooking-C} = T_{trans} + T_{prop} + 2T_{proc} \quad \text{Equation 6-4}$$

6.5.3 RWA Algorithm in Pre-booking Mechanism

During pre-booking, the RWA algorithm produces the decision for each request with a single

attempt. The result is either a rejection or a lightpath reservation with the appropriate edge delay, and the algorithm only reserves lightpaths during the transmission and propagation period.

More precisely, the algorithm is an extension of a traditional RWA, such as AUR-E [DZB04], and it can be illustrated as shown in Fig 6-5.

Firstly, the algorithm looks up the current lightpath reservation states to find all the wavelengths and links available at time $t_{earliest}$, and constructs w graphs (where w is the number of wavelengths) for later calculation.

Secondly, the RWA procedure runs the Dijkstra algorithm on each graph, and selects the shortest path as a candidate lightpath.

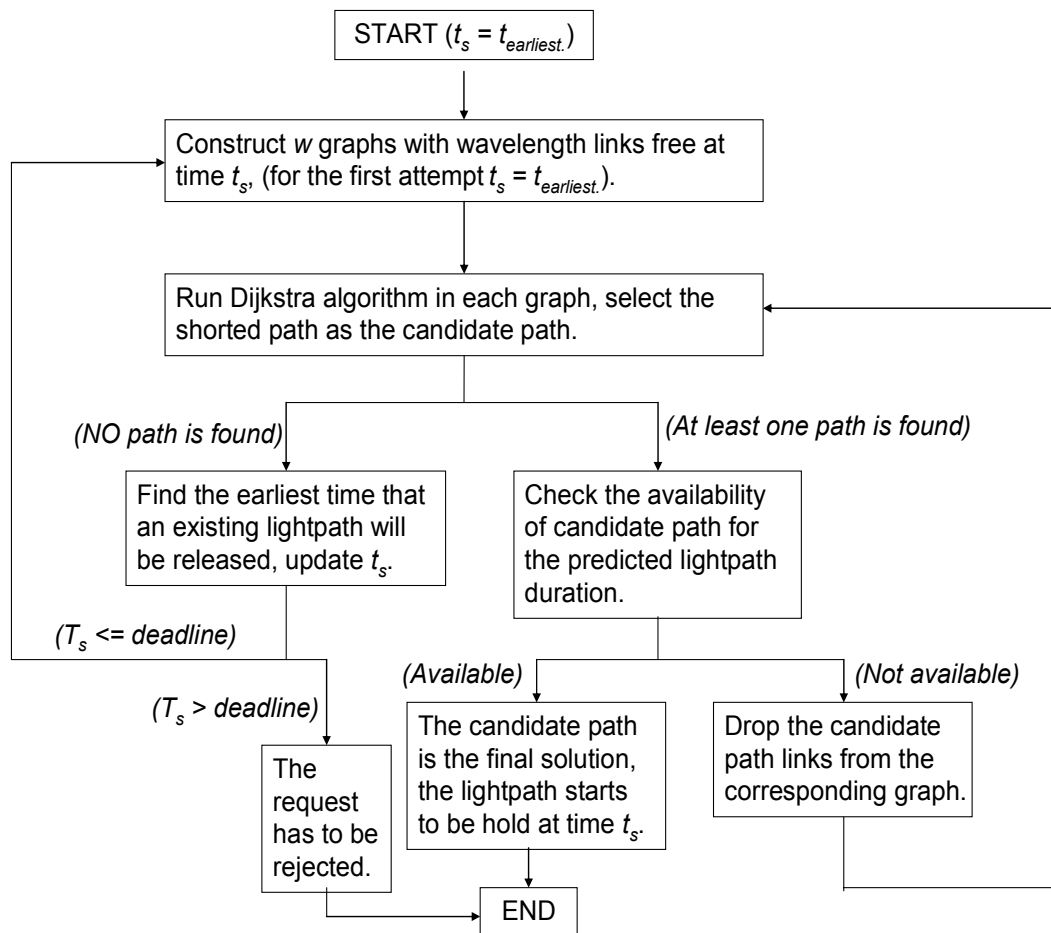


Fig 6-5: RWA algorithm in pre-booking mechanism

Thirdly, the selected lightpath availability is evaluated over the whole predicted lightpath duration. If the lightpath is fully available for this duration, the procedure stops and provides the resulting lightpath as the final solution. However, if the lightpath is not available for the whole

of the requested lightpath holding time, the procedure drops those wavelengths and links from the graph, and runs Dijkstra algorithm again based on the reduced graph. The Dijkstra algorithm will be repeated until the lightpath is found or no links are available in the graphs.

If no lightpath is found at time $t_{earliest}$, the RWA procedure will then find the earliest time that an existing lightpath will be released, and repeat the graph construction and Dijkstra algorithm to find a lightpath starting at the delayed time point.

In the end, if the delay exceeds the maximum delay that the traffic can tolerate, the request will be rejected.

Although the RWA algorithm described above is more computationally complicated than a traditional RWA, it allows the reservation of wavelength resources at future times with only one request attempt. Conversely, a traditional RWA can only find the lightpath availability at the time it is executed, so a traditional RWA has to be run multiple times at later time points to find out whether the lightpath is available at the specified time.

6.5.4 Comparison with WR-OBS

In summary, there are three differences between the pre-booking approach and the WR-OBS mechanism.

- (1) In pre-booking, the reservation request is initialized at the start of burst aggregation, whilst, in WR-OBS the lightpath request starts after an initial aggregation time. Given the same end-to-end delay requirement, the lightpath provisioning in pre-booking can be delayed longer than the WR-OBS, thus increasing the chance of reattempting a lightpath.
- (2) Pre-booking derives a lightpath solution before the transmission starts, and the lightpath is only held for the payload transmission and propagation period. Conversely, WR-OBS reactively reserves lightpaths, and the lightpath starts to be held during lightpath acknowledgement phase.
- (3) In pre-booking, the RWA algorithm calculates a delayed lightpath for the request with a single attempt. In WR-OBS, the RWA only checks the lightpath availability at current time, and the request has to be further queued for re-consideration if the current RWA attempt fails to find a suitable lightpath.

6.6 Implicit Pre-booking Signalling with Distributed Control

In a distributed control plane, the signalling and the routing decisions are independently made by individual nodes. Typically, with a MPLS/GMPLS-based control plane, the wavelength-based Label Switched Path (LSP) can be established and withdrawn by the ReSerVation Protocol (RSVP) or the Constrained Routing Label - Distribution Protocol (CR-LDP) [RFC3031][RFC3473].

6.6.1 Traditional Signalling and Reservation

Fig 6-6 illustrates an example of traditional LSP reservation with one intermediate core node [SST03].

The procedure starts with a Path message being sent from the ingress towards the egress. The Path message decides the route, collects possible labels, and establishes protocol state along the way.

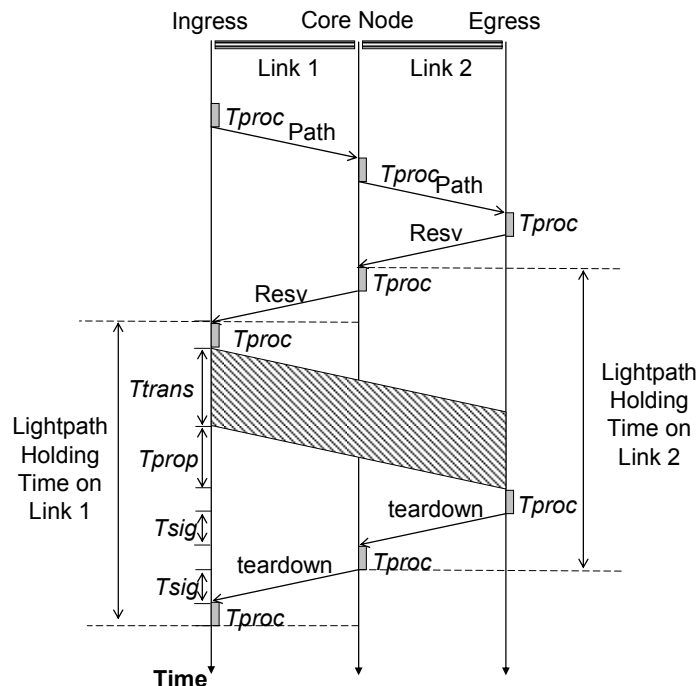


Fig 6-6: Traditional LSP reservation in distributed control plane

Then, a Resv message is sent in the reverse direction towards the ingress node, finalising the protocol state and configuring the OXC to start the resource holding at each LSP node. When the Resv reaches the ingress, the LSP is established, and the burst transmission starts.

The LSP tear down can be initiated from the egress node when the burst transmission completes⁴⁶. A teardown message is sent towards the ingress to release the OXC configuration and remove the protocol state along the way.

As illustrated, the lightpath holding time on each LSP link starts from the arrival of the Path message until the receipt of teardown message.

Assuming that the burst transmission time is T_{trans} , the end-to-end burst propagation time is T_{prop} , any signalling processing time at each node is T_{proc} , and the signal propagation time on each link is T_{sig} , then, the lightpath holding time of traditional LSP reservation can be expressed in Equation 6-5, where n is the total number of the lightpath nodes (including the ingress and the egress), and the factor $(n-1)T_{sig}$ is actually the total end-to-end signalling propagation time.

$$T_{lightpath}^{tradition-D} = T_{trans} + T_{prop} + (n+1)T_{proc} + (n-1)T_{sig} \quad \text{Equation 6-5}$$

6.6.2 Proposed Signalling and Reservation

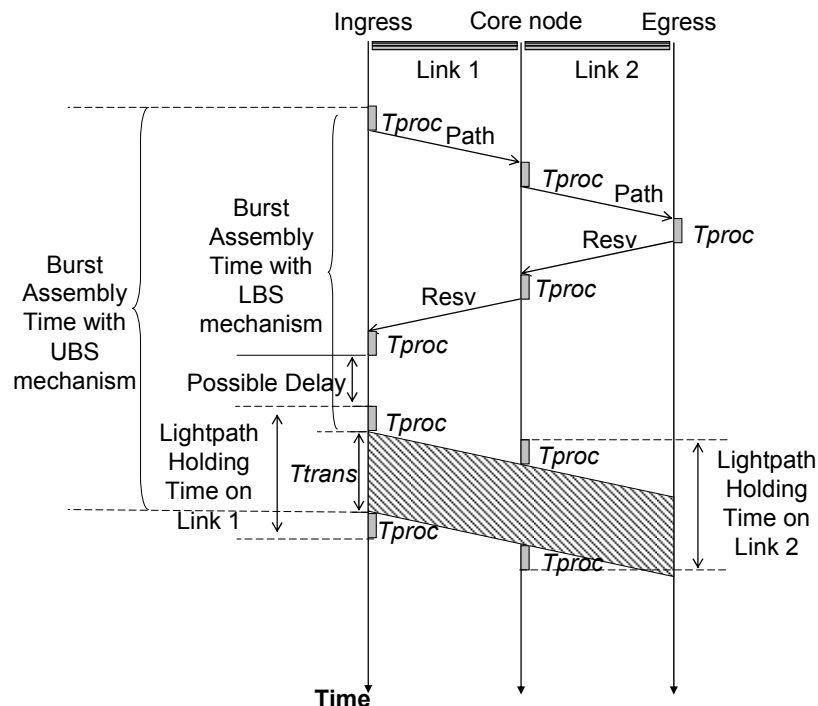


Fig 6-7: Pre-booking LSP reservation in distributed control plane

⁴⁶ The LSP tear down can also be initiated from the ingress node.

Correspondingly, Fig 6-7 presents an example of our pre-booking reservation using a distributed control plane.

Once the first bit of the new burst arrives at the assembly queue, the pre-booking starts with a Path message. In the Path message, the associated burst information, such as the earliest burst transmission time, the delay tolerance, and the burst length statistics are carried to help find a free label during the burst transmission window at each LSP node.

Then, the *Resv* message traverses in the reverse direction to confirm the reservation and schedule the OXC configuration at the pre-booked time points. While the OXC is not immediately setup on the arrival of the *Resv*, it is controlled by the software scheduling based on the committed reservation.

The resource release in pre-booking is not triggered by the teardown message. Instead, it is scheduled based on the burst length prediction.

The lightpath holding time of our pre-booking mechanism mainly consists of the burst transmission time and the OXC configuration time, and it can be expressed in Equation 6-6.

$$T_{lightpath}^{prebooking-D} = T_{trans} + 2T_{proc} \quad \text{Equation 6-6}$$

Therefore, pre-booking saves the majority of resources during the signalling propagation time and the burst propagation time. In a large backbone with a very high wavelength transmission rate, the burst transmission time will be very short, whilst the propagation time remains in the tens of millisecond regime. The resource saving from the signalling propagation time is significant, and the saved resources can be used to support more traffic.

6.7 Analysis of Lightpath Bandwidth Efficiency

The potential benefit of our approach can be expressed as the lightpath bandwidth efficiency. The higher is the lightpath bandwidth efficiency, the more efficiently the resource is utilized.

6.7.1 Centralized Schemes

The lightpath holding time in the centralized control plane has been described in section 6.5. More specifically, the signalling propagation time between the edge node and the central node

can be approximated as half of the burst end-to-end propagation time, that is $T_{sig} = 0.5T_{prop}$. Thus, the lightpath bandwidth efficiency for the traditional reservation mechanism, WR-OBS, and the pre-booking mechanism, in the centralized control plane, can be expressed as in Equation 6-7, 6-8, and 6-9, respectively.

$$E^{tradition-C} = \frac{T_{trans}}{T_{trans} + T_{prop} + 3 \times 0.5T_{prop} + 3T_{proc}} \quad \text{Equation 6-7}$$

$$E^{WROBS} = \frac{T_{trans}}{T_{trans} + T_{prop} + 0.5T_{prop} + T_{proc}} \quad \text{Equation 6-8}$$

$$E^{prebooking-C} = \frac{T_{trans}}{T_{trans} + T_{prop} + 2T_{proc}} \quad \text{Equation 6-9}$$

Let the OXC processing time T_{proc} be 1ms, the lightpath bandwidth efficiency can be deployed as shown in Fig 6-8.

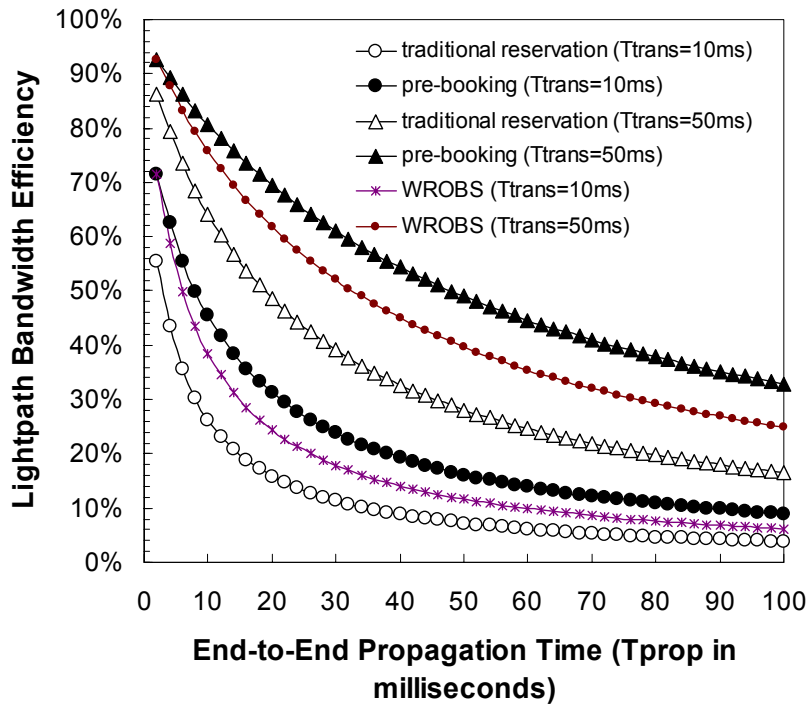


Fig 6-8: Lightpath bandwidth efficiency in centralized control plane

Fig 6-8 clearly illustrates that the lightpath bandwidth efficiency is directly affected by the reservation protocols. With the centralized control plane, the traditional reservation yields the lowest lightpath bandwidth efficiency, as the lightpath is held for the longest time among the

three schemes. However, some of the resources in the traditional reservation are held whilst idle, and actually the lightpath does not need to be held for so long. The lightpath bandwidth efficiency for WR-OBS is between the traditional reservation and the pre-booking mechanisms, because the WR-OBS assumes that the lightpath is promptly released once the transmission completes, whilst some resources during the lightpath acknowledgement phase are still unnecessarily held. The pre-booking mechanism only consumes the resources during the burst transmission and propagation, excluding any wastage during the signalling stages, so it achieves the highest lightpath bandwidth efficiency. The lightpath bandwidth efficiency in the pre-booking mechanism decreases with the increase of propagation time, because the lightpath holding time in this circumstance also includes the burst propagation time⁴⁷. Additionally, the advantage of our pre-booking mechanism becomes significant when the burst transmission time is of a magnitude comparable to the end-to-end propagation delay.

6.7.2 Distributed Schemes

With the distributed control plane, the burst end-to-end propagation time usually equals to the signalling end-to-end propagation time (excluding the signal processing time at each OXC), so in Equation 6-5, the factor of $(n-1)T_{sig}$ can be approximated as T_{prop} . As a result, the lightpath bandwidth efficiency for the traditional reservation mechanism and the pre-booking approach in the distributed control plane can be expressed as in Equation 6-10 and 6-11, respectively (where n is the number of nodes along the lightpath).

$$E^{tradition-D} = \frac{T_{trans}}{T_{trans} + T_{prop} + (n+1)T_{proc} + T_{prop}} \quad \text{Equation 6-10}$$

$$E^{prebooking-D} = \frac{T_{trans}}{T_{trans} + 2T_{prop}} \quad \text{Equation 6-11}$$

Let $n = 4$ (the lightpath consists of four nodes, three links), and let the OXC processing time T_{proc} is 1ms, the lightpath bandwidth efficiency can be illustrated in Fig 6-9. Fig 6-9 demonstrates that, with the distributed control plane, the pre-booking mechanism yields a constant lightpath bandwidth efficiency. This is because the lightpath holding only includes the

⁴⁷ Including the burst propagation time in the lightpath holding time allows the RWA algorithm to remain simple.

burst transmission time and the OXC processing time, and pre-booking actually achieves the maximum lightpath bandwidth efficiency. However, traditional reservation yields much lower lightpath bandwidth efficiency than the pre-booking mechanism. This is caused by unnecessary resource holding during the signalling propagation and the burst propagation time. Additionally, the gain of pre-booking over traditional reservation increases with the end-to-end propagation time, so our pre-booking scheme provides notably higher lightpath bandwidth efficiency than the traditional reservation mechanisms when applied to large-scale backbones.

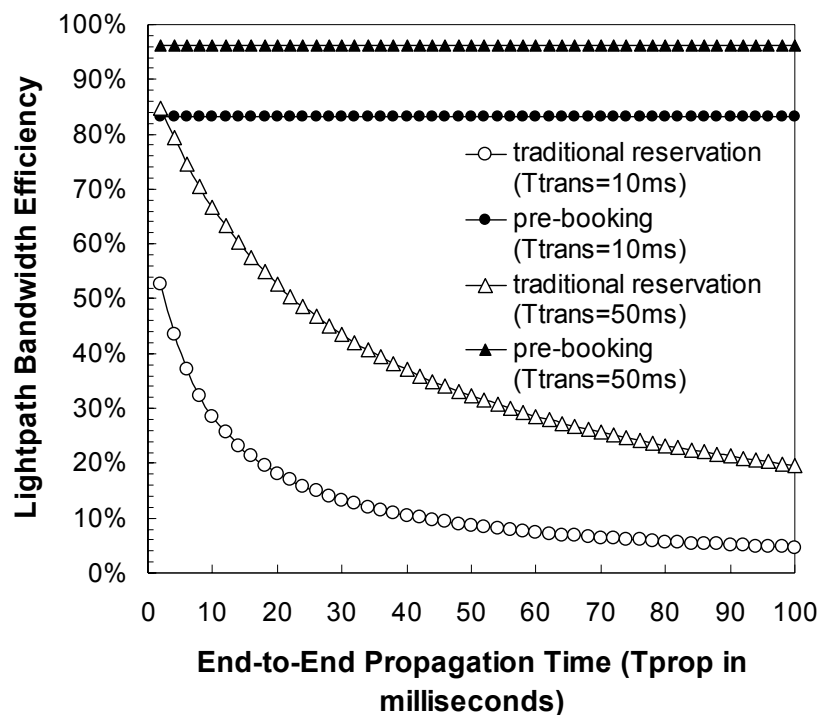


Fig 6-9: Lightpath bandwidth efficiency in distributed control plane

The analysis of lightpath bandwidth efficiency shows that the pre-booking mechanism achieves the highest bandwidth efficiency in both the centralized and the distributed control planes. The enhanced bandwidth efficiency demonstrates the potential of the pre-booking mechanism and it is expected to significantly improve network performance and decrease the traffic loss in a WAN/backbone environment.

6.8 Burst Length Study with Self-Similar Incoming Traffic

In pre-booking, the burst length information is required to indicate the resource releasing time. The burst length estimation should be based on traffic statistics. Current network traffic

observations identify self-similar behaviour [EA97][ESW94][ZS99][AMD02], so our work focuses on studying burst length with self-similar incoming traffic.

6.8.1 Definition of Self-Similarity

Let $X = (X_t; t = 0, 1, 2, 3, \dots)$ be a covariance stationary stochastic process with mean μ , variance σ^2 , and autocorrelation function $r(k)$, $k \geq 0$. In particular, we assume that the autocorrelation function $r(k)$ has the form of Equation 6-12.

$$r(k) \sim k^{-\beta} L(k) \quad \text{Equation 6-12}$$

Where $0 < \beta < 1$, and L is slowly varying at infinity (for simplicity, L can be assumed as asymptotically constant).

To explain the self-similarity, another covariance stationary stochastic process $X^{(m)} = (X_i^{(m)}; i = 0, 1, 2, 3, \dots)$ with the corresponding autocorrelation function $r^{(m)}(k)$, has to be introduced. The $X^{(m)}$ is derived by averaging the original series X over non-overlapping blocks of size m . That is, $X_i^{(m)} = \frac{1}{m}(X_{im-m+1} + \dots + X_{im}), i \geq 1$.

The process X is said to be *exact second-order self-similar* with the Hurst parameter $H = 1 - \beta/2$, if for all $m = 1, 2, 3, \dots$, the variance and the autocorrelation function of the process $X^{(m)}$ satisfies Equation 6-13 and Equation 6-14.

$$\text{var}(X^{(m)}) = \sigma^2 \times m^{-\beta} \quad \text{Equation 6-13}$$

$$r^{(m)}(k) = r(k) \quad (k > 0) \quad \text{Equation 6-14}$$

The process X is called *asymptotically second-order self-similar* with the Hurst parameter $H = 1 - \beta/2$, if for all k large enough, the autocorrelation function of process $X^{(m)}$ satisfies Equation 6-15.

$$r^{(m)}(k) \rightarrow r(k) \quad \text{as } m \rightarrow \infty \quad \text{Equation 6-15}$$

As a conclusion, exact or asymptotically second-order self-similar process X yields indistinguishable aggregated process $X^{(m)}$, at least with respect to their autocorrelation

functions. More intuitively, the shape of process $X^{(m)}$ is very similar to that of X , and the aggregation will not smooth the burstiness of the original series X .

6.8.2 Scaling Behaviours in Internet Traffic

The Internet traffic behaviour is affected by multiple factors like the TCP transport protocol and the file size distribution. Traffic collected from LANs exhibits complicated scaling behaviours [GW99][AMD02]. Generally speaking, there is a cutoff timescale beyond which the traffic exhibits mono-fractal scaling (self-similar) with a constant Hurst parameter H , whilst within short timescales the traffic displays complex multi-fractal behaviour.

Mono-fractal scaling means *H-self-similarity*, and the degree of self-similarity can be measured by a fixed *Hurst* parameter. Multi-fractal means that the degree of self-similarity is not constant. Instead, the *Hurst* parameter is a time-varying variable, which can be presented as $H(t)$ [CCL97]. As mono-fractal traffic and multi-fractal traffic yield different statistic characteristics, and require different models to generate the corresponding traffic traces for simulation study [ESW94][DW99], it is important to define whether the traffic under study is mono-fractal or multi-fractal.

In LANs, the cutoff timescale is around the round-trip time on the order of a few hundred milliseconds or seconds. The reason for the characteristic to be mono-fractal behaviour over large timescale is mainly due to the heavy-tailed distribution of file sizes over the Internet, whilst, the multi-fractal behaviour within small timescales is mainly due to the protocol dynamics such as TCP flow control, network congestion, packet loss and packet retransmission [GW99].

However, in ultra-high speed long-haul WDM backbones (beyond 10Gbps), the cutoff timescale is expected to decrease due to the high aggregation level [IA02][YPS05]. Moreover, it has been observed that the traffic trace collected from OC3/12/48 links of a US tier-1 backbone network manifests a mono-fractal behaviour within timescales from 1 to 100 milliseconds, and the aggregated traffic is Gaussian distributed [ZRM03][CFC03][YPS05].

In the pre-booking scheme, a large number of users send traffic to the SPs at the backbone edge, and the bursts/flows aggregation time ranges from milliseconds to tens of milliseconds, so we

can assume that the traffic within the aggregation timescale is stationary H -self-similar and mono-fractal. Additionally, all the OBS research to-date assumes the traffic within the aggregation time scale to be simple H -self-similar [XVC00][DB02][IA02][ZRM03][CFC03].

6.8.3 Burst Length Statistics

With H -self-similar incoming traffic, and the burst assembly procedure at the network edge, there is a wide agreement among many researchers that, if the aggregation time is much longer than the packet length (in time scale), *Central Limit Theorem* (CLT) applies, and the burst length follows a normal distribution [XVC00][DB02][IA02][ZRM03][CFC03].

More specifically, by the definition of self-similarity and ref [IA02], the burst length can be quantified as a Gaussian random variable with mean value B expressed in Equation 6-16, and the variance ν as shown in Equation 6-17, where T_a is the burst aggregation time, μ is the average incoming traffic bit rate, σ is its standard deviation, and H is the Hurst parameter.

$$B = T_a \times \mu \quad \text{Equation 6-16}$$

$$\nu = T_a^{2H} \times \sigma^2 \quad \text{Equation 6-17}$$

Equation 6-16 and Equation 6-17 indicate that the burst length and related deviation are predictable as long as the incoming traffic bit rate and variance are measured correctly. Therefore, in the pre-booking mechanism, the average incoming bit rate and variance are closely monitored by the SPs, such that the burst length can be estimated based on Equation 6-16 and Equation 6-17.

6.9 Reservation Strategy based on Burst Length Prediction

Since the burst length follows a normal distribution with mean and variance as described in Equation 6-16 and 6-17, the burst length can be predicted by monitoring the recent incoming traffic parameters⁴⁸, μ , ν , and H .

⁴⁸ We assume that the average incoming traffic bit rate remains in a stable state for a long time. However, the incoming bit rate may fluctuate with time, so the observation may be inaccurate at the boundary when a fluctuation happens, and a short period of time is required for the observations to return to an accurate level. In the case when

More specifically, the burst length is estimated as the mean value plus k times the corresponding standard deviation, where k is an adjustable parameter, and the predicted burst length is expressed in Equation 6-18.

$$L_{predicted} = T_a \times \mu + k \times T_a^H \times \sigma \quad \text{Equation 6-18}$$

In Equation 6-18, $T_a \times \mu$ represents the average burst length according to Equation 6-16, and $T_a^H \times \sigma$ corresponds to the burst length standard deviation according to Equation 6-17.

We introduce extra k times of standard deviation in the total reserved burst duration, such as to decrease the probability that the actual burst length is longer than the reservation time window. Although this causes some resource wastage, in future WDM backbones, the wavelength transmission rate can be as high as several $Tbps$, the time needed for the extra bits transmission is very short, and the pre-booking still achieves resource savings from the signalling propagation time, which is in the millisecond region (about 1ms per 200km fibre). We call our reservation strategy “excess reservation”, and expect it to effectively avoid high bit-loss rates caused by reservation inaccuracies.

As the burst length estimation is related with the burst aggregation time, the following two sub-sections describe the burst length prediction associated with LBS and UBS burst assembly mechanisms.

6.9.1 Burst Length Reservation Strategy with LBS Assembly Mechanism

In the LBS assembly mechanism, the burst aggregation starts when the first bit of the burst arrives at the assembly queue, denoted as t_s . The burst aggregation ceases when the acknowledgement is received by the edge node, denoted as t_{ack} . The total burst aggregation time is $T_a = t_{ack} - t_s$. Therefore, the reserved burst length based on estimation can be expressed as in Equation 6-19.

$$L_{predicted}^{LBS} = (t_{ack} - t_s) \times \mu + k \times (t_{ack} - t_s)^H \times \sigma \quad \text{Equation 6-19}$$

the traffic varies very frequently, a time-series model, such as the LMS-based LPF (refer to Section 3.2.1), can be used to predict the bit rate.

6.9.2 Burst Length Reservation Strategy with UBS Assembly Mechanism

In the UBS assembly mechanism, the burst aggregation starts at time t_s , when the first bit of new burst arrives. The burst aggregation continues until the burst starts transmission and the buffer is emptied. Assuming that the existing burst length when the burst starts transmission is L_1 , the time required to empty the buffer, t , can be calculated from Equation 6-20, where $l(t)$ is the extra incoming burst length after the burst starts transmission, C is the outgoing wavelength rate.

$$L_1 + l(t) = C \times t \quad \text{Equation 6-20}$$

According to our excess reservation strategy as shown in Equation 6-18, and letting the time that the burst transmission starts be t_{trans} , L_1 can be estimated as in Equation 6-21, and $l(t)$ can be expressed as in Equation 6-22.

$$L_1 = (t_{trans} - t_s) \times \mu + k \times (t_{trans} - t_s)^H \times \sigma \quad \text{Equation 6-21}$$

$$l(t) = t \times \mu + k \times t^H \times \sigma \quad \text{Equation 6-22}$$

If we ignore the $k \times t^H \times \sigma$ factor in Equation 6-22, the time required to empty the buffer, t , can be deduced by substituting Equation 6-22 with Equation 6-20 and 6-21, and it can be expressed as shown in Equation 6-23.

$$t = \frac{L_1}{C - \mu} = \frac{(t_{trans} - t_s) \times \mu + k \times (t_{trans} - t_s)^H \times \sigma}{C - \mu} \quad \text{Equation 6-23}$$

The total burst length is actually the length being transmitted during time t , so the estimated burst length can be expressed as described in Equation 6-24.

$$L_{predicted}^{UBS} = C \times t = \frac{C}{C - \mu} \times [(t_{trans} - t_s) \times \mu + k \times (t_{trans} - t_s)^H \times \sigma] \quad \text{Equation 6-24}$$

6.10 Bit Loss Rate Caused by Insufficient Reservation

In proposed pre-booking mechanism, burst length prediction is required to determine the resource reservation duration. However, prediction can be inaccurate. Therefore, we adopt the

following strategy to make efficient use of reserved resources whilst coping with the prediction error.

More specifically, if the predicted burst length is longer than the real burst length, the whole burst will be successfully delivered, though some of communication resources are wasted. Conversely, if predicted burst length is shorter than the actual burst length, given the big volume of the burst, instead of abandoning the whole burst, only the residual portion is dropped. Thus, bit loss happens when the predicted burst length is shorter than the actual burst length, and the reservation is insufficient.

Focusing on the traffic loss caused by prediction error and insufficient pre-booking, we can assume that the burst will never be blocked by the core network. Bit loss rate caused by insufficient reservation can be calculated as in Equation 6-25.

$$\begin{aligned}
 \text{bit_loss_rate} &= \frac{\sum \text{lost_bits}}{\sum \text{burst_bits}} \\
 &= \frac{\sum_t P(l(t) > L_p) \times (l(t) - L_p)}{\sum_t l(t)} \\
 &= \frac{\int_{L_p}^{\infty} (x - L_p) f(x) dx}{\int_0^{\infty} x f(x) dx} \\
 &= \frac{\frac{\sigma}{\sqrt{2\pi}} \times \exp\left(-\frac{(L_p - \mu)^2}{2\sigma^2}\right) + (\mu - L_p) \times N\left(\frac{L_p - \mu}{\sigma}\right)}{\frac{\sigma}{\sqrt{2\pi}} \times \exp\left(-\frac{\mu^2}{2\sigma^2}\right) + \mu \times N\left(-\frac{\mu}{\sigma}\right)}
 \end{aligned}
 \tag{Equation 6-25}$$

Where $l(t)$ is the burst length (in bits) at time t , L_p is the predicted burst length, $P(l(t) > L_p)$ is the probability that the real burst length longer than the predicted burst length, x is a general term for burst length, $f(x)$ is the PDF of burst length following normal distribution with mean μ and standard deviation σ , and $N(\bullet)$ is the Q-Function of the standard normal distribution. The results for bit loss rate caused by insufficient pre-booking will be further demonstrated with together with the simulation results in next chapter.

6.11 Summary

This chapter describes the pre-booking mechanism for connectivity acknowledged OBS networks. The network architecture and the signalling procedure are described for both a centralized control plane and the distributed control plane. The analysis for lightpath bandwidth efficiency outlines the prime benefits achieved by the pre-booking mechanism.

Burst length prediction is important in the pre-booking mechanism, because it helps to decide when to release the network resource without waiting for the receipt of a Release message. With widely observed self-similar traffic and burst aggregation at the network edge, this research accepts the common premise that the aggregate burst length follows a normal distribution. Accordingly, the research proposes an excess reservation strategy, where the reservation duration is the mean burst duration plus k times its standard deviation, such that the probability that the real burst length is larger than the reserved length is reduced and less bits are lost due to insufficient reservations. Furthermore, the bit loss rate due to the insufficient reservations are mathematically analysed and deduced in section 6-10.

Chapter 7 Simulation Models and Results for Pre-booking Mechanism in Connectivity Acknowledged OBS Networks

This chapter considers the centralized control plane as an example, and constructs simulations via OPNET™ modeller to demonstrate the benefits of the pre-booking mechanism.

There are mainly two reasons to choose the centralized control plane for the simulation. Firstly, there are no previously published results to demonstrate the OBS performance with the MPLS/GMPLS signalling protocol, whilst, there are many results published for WR-OBS with a centralised control plane. Therefore, using centralized control plane enables a performance comparison with one of the existing schemes. Secondly, it is worth mentioning that the centralized control plane does not over-simplify the simulation scenarios. The centralized control plane provides global lightpath control for all the requests, and this approach allows the maximum flexibility in routing and wavelength selection based on real-time updates of the global link-state database. In the distributed approach, the reservation is done on a hop-by-hop basis, where each node maintains a link-state database, a routing table and per-link reservation tables. The update of the routing table is not instant, so there can be bottleneck nodes due to the lack of up-to-date comprehensive network information. Nevertheless, choosing the centralized control plane to demonstrate the advantages of pre-booking mechanism is still appropriate, as the pre-booking mechanism achieves performance gain by excluding unnecessary resource holding at the signalling propagation stage. Routing is a totally different and independent issue that will affect the network performance, so this study can eliminate the routing affects, and use the centralised control plane to reveal the performance gain obtained by the pre-booking mechanism.

7.1 Simulation Models

Within the simulation, there are basically three types of models. The first is the traffic model. For Constant Bit Rate (CBR) traffic, the traffic intensity can be simply defined by the incoming bit rate, and the burst length can be calculated by the burst aggregation time. For self-similar

traffic, there are multiple approaches, and they will be described in a later section. The second type of simulation model is the central control node, which maintains network link-state database and deals with lightpath requests. The third type of model is the OBS edge node, which aggregates bursts, manipulates requests, and sends/drops bursts at the network edge.

The OBS core node is not modelled in the simulation, because the functionality of the core nodes is simply relaying bursts to the next node. More specifically, in centralized control plane, the central node contains a complete knowledge of the whole network. The activity of the core optical node is totally controlled by the central node, so the blocking probability and the lightpath bandwidth efficiency are actually determined by the decision of the central node. Once the lightpath is committed by the central node, the burst is transparently delivered from the source to the destination, so the burst transmission can be virtually assumed by the central node, and there is no need to model the core node in the simulation. The elimination of core nodes in the simulation can reduce the number of simulation events without affecting the results, thus effectively accelerating the simulation.

In addition, for comparison purposes, three schemes have been implemented. They are the pre-booking mechanism, the WR-OBS mechanism, and a traditional reservation mechanism. The WR-OBS mechanism is different from the traditional reservation mechanism in that the former scheme releases the network resources once the last bit of the burst reaches the destination node⁴⁹, whilst, the latter scheme holds the resource until the RELEASE message is received.

7.1.1 H-Self-Similar Traffic Modelling

A widely adopted method for generating H-self-similar (mono-fractal) traffic is the superposition of a number of ON-OFF models, with either ON or OFF, or both periods following heavy-tailed distribution [ESW94][EA97][DB02].

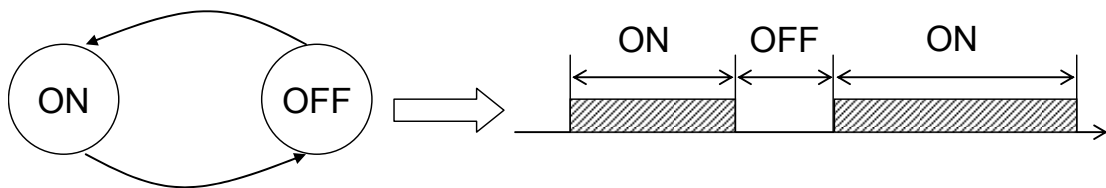
⁴⁹ Actually, WR-OBS assumes ideal resource releasing. In most of the existing WR-OBS research, perfect knowledge of burst length or constant bit rate traffic is assumed such that the lightpath can be torn down at the appropriate time.

7.1.1.1 Superposition of ON-OFF Models

Self-similar traffic can be constructed by superimposing many simple renewal reward processes, in which the reward is restricted to the values 0 and 1, and the renewal intervals are heavy-tailed [EA97].

7.1.1.1.1 ON-OFF Process

The renewal reward process can be represented as an ON-OFF process, as shown in Fig 7-1.



ON and OFF periods are independent random processes. Either or both of them should yield a heavy-tailed distribution.

Fig 7-1: The ON-OFF Process

The ON-OFF process strictly alternates between the ON and OFF periods, where ON and OFF periods are separate independent random processes. During the ON period, the reward yields the value 1, indicating that the packet or traffic bits are in the transmission phase. During the OFF period, the reward yields the value 0, indicating the idle times between transmissions.

Additionally, to ensure LRD, either the ON period should follow a heavy-tailed distribution with parameter α_1 , or the OFF period follows a heavy-tailed distribution with parameter α_2 , or both.

If there are a large number of such ON-OFF processes independently and concurrently in operation, the aggregation of these sources will be a self-similar process, with $H = (3 - \min(\alpha_1, \alpha_2)) / 2$ [EA97]. More strictly speaking, traffic generated by an infinite number of ON-OFF sources is strictly second-order self-similar. However, the traffic from a finite number of ON-OFF sources is asymptotic second-order self-similar [EA97][WPT98].

7.1.1.1.2 Heavy-Tailed Distribution

A distribution is said to have a heavy-tail if the probability of large values occurring is non-negligible, that is, Equation 7-1.

$$P(X > x) \sim x^{-\alpha}, \text{ as } x \rightarrow \infty, 0 < \alpha < 2 \quad \text{Equation 7-1}$$

The simplest heavy-tailed distribution is the Pareto distribution, in which the PDF function follows Equation 7-2.

$$f(x)_{\text{pareto}} = \frac{\alpha k^\alpha}{x^{\alpha+1}}, \quad k \leq x < \infty \quad \text{Equation 7-2}$$

In the Pareto distribution (Equation 7-2), $k > 0$ is the location parameter representing the smallest possible value of the random process, and $\alpha > 0$ is the shape parameter controlling the shape of the distribution. When $\alpha \leq 2$, the random process yields infinite variance. Additionally, when $0 < \alpha \leq 1$, the process has an infinite mean, whilst, when $\alpha > 1$, the mean value of the random process can be expressed as in Equation 7-3.

$$\mu_{\text{pareto}} = \frac{\alpha k}{\alpha - 1} \quad (\alpha > 1) \quad \text{Equation 7-3}$$

To generate a LRD source with a stable mean value, we should use a Pareto distribution with $1 < \alpha < 2$.

7.1.1.1.3 The Self-Similar Source Model with Specified Incoming Intensity

The model used in our simulation for generating self-similar traffic can be illustrated by Fig 7-2.

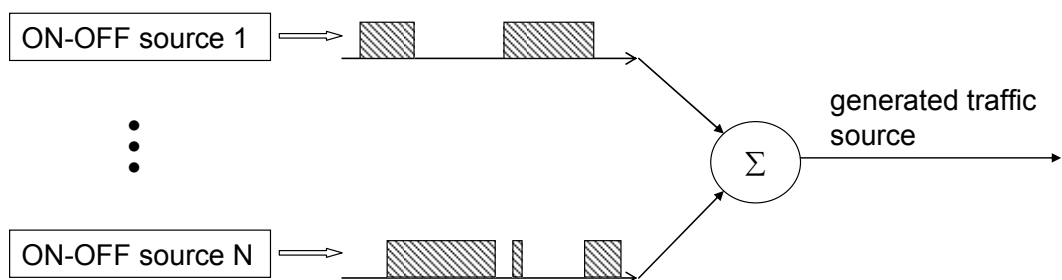


Fig 7-2: Superposition of ON-OFF processes

N ON-OFF processes are provided to generate the self-similar traffic. Each On-OFF process transmits traffic with a bit rate b_{in} , and all the ON-OFF processes yield the same parameters, though they are independent to each other. The resulting traffic will have a total incoming bit rate, B_{in} , which represents the incoming traffic intensity, and is expressed as in Equation 7-4.

$$B_{in} = \frac{E(ON) \times b_{in}}{E(ON) + E(OFF)} \times N \quad \text{Equation 7-4}$$

In Equation 7-4, $E(ON)$ and $E(OFF)$ are the average ON and OFF periods, respectively, and they can be calculated from Equation 7-3. If we fix the parameters for ON and OFF periods, and we use a fixed number of sources, the resulting incoming bit rate, B_{in} , can be easily adapted by the individual incoming bit rate, b_{in} . Therefore, such a model is very easy to control in the simulation.

Moreover, in packet networks, traffic is usually represented by packets of various sizes [XVC00]. However, in our simulation model, we do not packetize the traffic. Instead, we assume that the ON period is fully occupied with traffic bits, and these traffic bits are transmitted with a constant bit rate b_{in} . This is partly due to the fact that we only interested in the burst loss rate and the bit loss rate in our study. Moreover, packetizing traffic requires extra specification for the packet size and inter-arrival time within the ON-period, and the last packet arriving within the given ON-period might have to finish its transmission outside the ON period, thus changing the self-similarity of the final generated traffic [XVC00].

7.1.1.2 Validating the Self-Similarity

The self-similarity of generated source can be validated by a variance-time plot. More specifically, according to the definition of a second-order self-similar process (as in Section 2.4.2.2.1), the variances of aggregated processes $X^{(m)}$, $m \geq 1$, should satisfy Equation 2-3. Therefore, in log-log plots, the variances decrease linearly against m with a slope equal to $2(H-1)$ [ESW94]. Therefore, a straight line with slope greater than -1 is indication of self-similarity, where $slope = 2(H-1)$. The corresponding evidence is provided in appendix F.

7.1.2 Central Node Model

In all the three schemes under study, the central node model accepts requests from edge nodes into a central request buffer. The requests will be scheduled based on the Earliest Deadline First (EDF) principle, where the deadline for each request is determined by its delay tolerance. When the request is processed, a lightpath decision with or without extra delay, or a rejection is made by the RWA algorithm, and the corresponding acknowledgement is sent back to the edge node informing of the result. As we only consider one class of traffic, only one request buffer is provided in the simulation.

7.1.2.1 Central Node Model for WR-OBS and Traditional Reservation Schemes

The central node model for WR-OBS and traditional reservation schemes are the same and is illustrated in Fig 7-3.

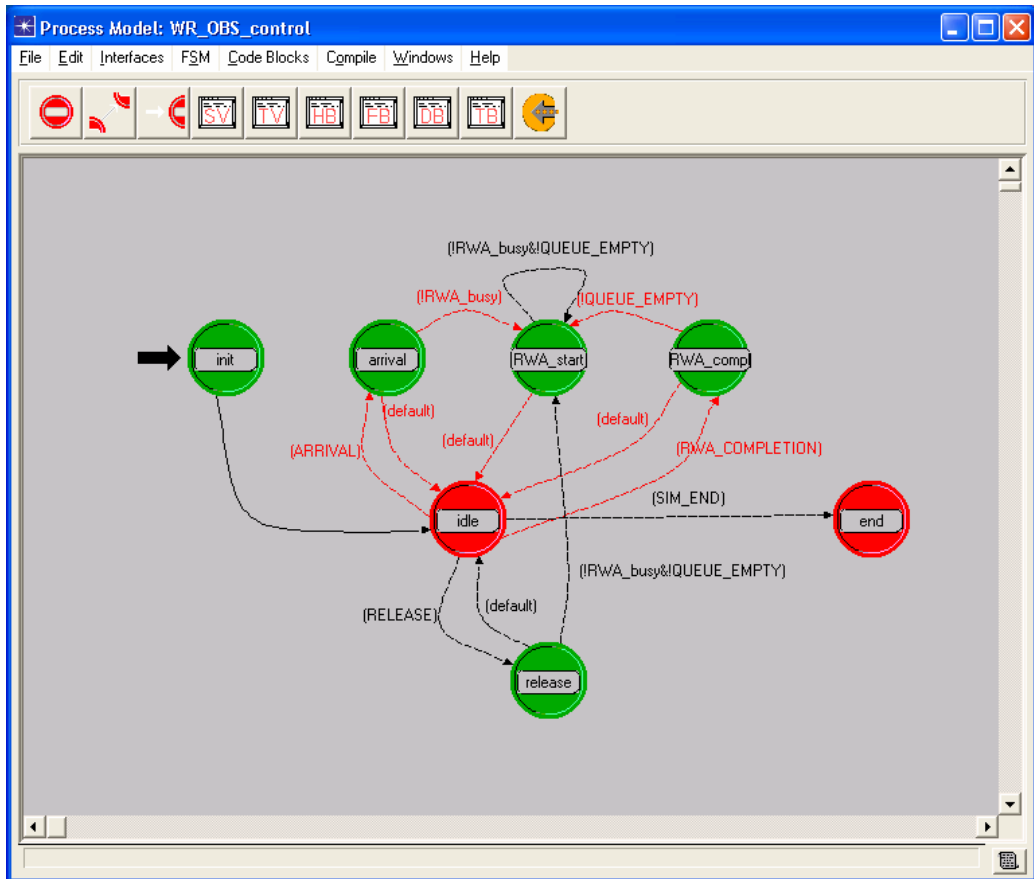


Fig 7-3: Central Node Process Model for WR-OBS and Traditional Reservation Schemes

The model consists of seven states. In the “init” state, the global link-state database and other parameters are initialized. The link-state database records the availability of each wavelength on each link in the network. At the beginning of the simulation, all the wavelengths and links are free for use. The “idle” state is an unforced temporary state, in which the simulation events are examined, and the simulation is transferred to a corresponding state.

When a request arrives at the central node, the simulation enters the “arrival” state, in which the request is queued into the central buffer. As each request carries information concerning the deadline, the position of the request in the buffer is determined based on EDF principle such

that the first request in the request buffer is always the one with the earliest deadline.

If the RWA processor is not busy with other requests and there are requests waiting in the request buffer, the simulation enters the “RWA_start” state. In this state, the first request in the request queue is extracted from the buffer. Additionally, the time for the current RWA calculation is given based on a random process following the Beta distribution, bounded by 0 and 0.2 ms, and with an average of 0.1ms [MKB02]. This RWA calculation time is the time required to run the RWA algorithm for the request.

The simulation then moves to the “RWA_compl” state after the expiry of RWA calculation time, indicating the end of RWA process. In this state, all the available wavelengths in corresponding links are used to construct graphs with specific wavelengths, and a lightpath search is carried out in each graph based on the Dijkstra’s algorithm. In case that one or more lightpaths are found by the RWA algorithm, the shortest lightpath is chosen as the lightpath offer to the request. The link-state database is updated, and all the involved wavelengths and links are set to occupied. Meanwhile, the acknowledgement is sent from the central node to the corresponding edge node, such that the burst can start transmission. However, if no lightpaths are available for the request in any of the graphs, the deadline of the request is examined to determine whether the request should be rejected or be further buffered for another attempt of the RWA process. More specifically, if the request’s delay tolerance is insufficient, the request will be rejected with a rejection message from the central node to the edge node. On the other hand, if the request can spend more time in the central node and try the RWA process at a later time without violating the delay tolerance, the request will be temporarily buffered in a delay buffer until some resources are released, and the request is considered by the RWA process again.

The “release” state is triggered by the release of any committed lightpaths. The global link-state database is updated, and the wavelengths and links along the released lightpaths are set to be free again. Moreover, if there are delayed requests in the delay buffer, the first request will be extracted and be considered by the RWA process to see if there is any lightpath available after the delay.

Finally, the “end” state finalises the simulation and records the simulation results.

7.1.2.2 Central Node Model for Pre-booking Scheme

The central node model for the pre-booking mechanism is similar to that of the WR-OBS and traditional reservation schemes, as shown in Fig 7-4.

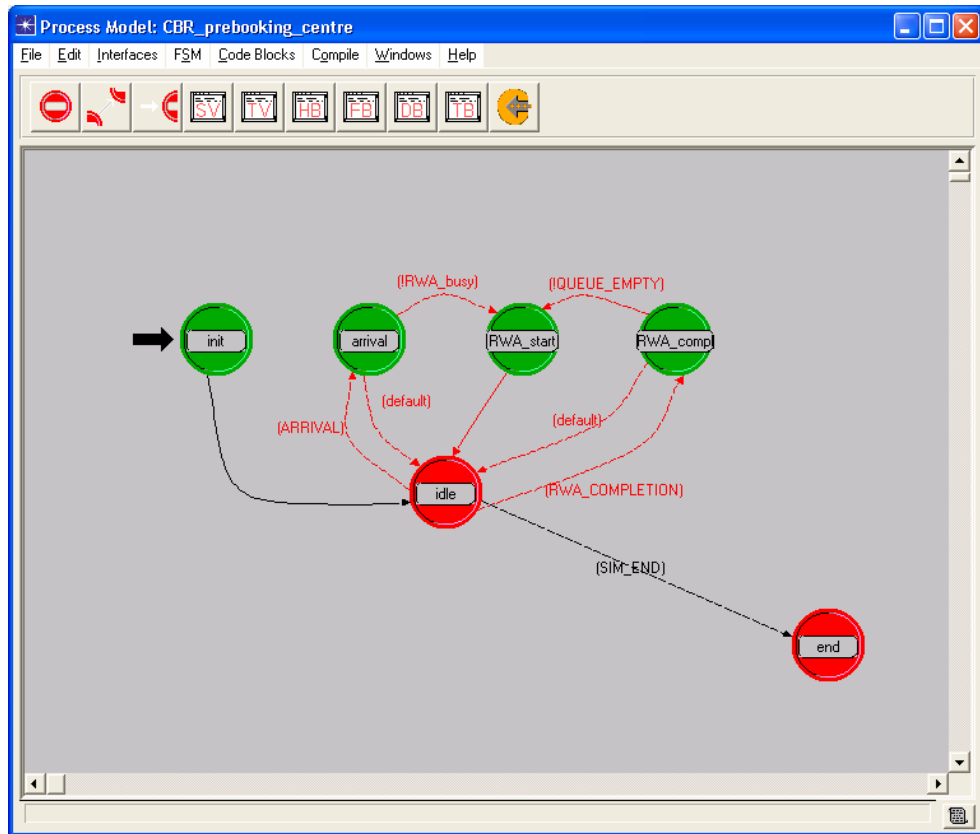


Fig 7-4: Central Node Process Model for Pre-booking Scheme

The main differences between the pre-booking central node and the other two schemes' central node are that the global link-state database in pre-booking scheme is a function of resources availability in relation to the simulation time, and the RWA algorithm in pre-booking scheme is more complicated than the one used in the other two schemes.

More specifically, in the pre-booking mechanism, resource reservation is made for a future time duration. The global link-state database has to record detailed information in terms of the wavelength and link availability at specific time points.

When the request is processed by the RWA algorithm, the RWA algorithm takes information such as the earliest burst transmission time and the delay tolerance as inputs. The details of the RWA algorithm has been described in Section 6.3.3, and most importantly, the adopted RWA algorithm is able to search for a lightpath at every possible time point based on the link-state

database and the delay tolerance specified by the request. Moreover, it is the central node that determines the time for burst transmission. Therefore, if the incoming traffic is self-similar, the burst length can be predicted by Equation 6-15 or Equation 6-20. As a consequence, the central node has the knowledge of the predicted burst length, and knows when to tear down the lightpath. The central node keeps all the information, and schedules the lightpath setup and release based on the committed reservation.

As the RWA algorithm is more complicated than the one used in the WR-OBS scheme, we set the RWA process to be time three times that of the WR-OBS case, and it is a random process following a Beta distribution, bounded by 0 and 0.6 ms, and with average 0.3ms.

7.1.3 Edge Node Model

The edge node model represents the SP's edge device with functionality for burst aggregation, lightpath requests, and burst transmission.

We assume that each node in the network is connected with an SP edge node, so we have N edge nodes in the network model, where N is the number of nodes in the topology. Moreover, as each SP edge node sends bursts to all the other nodes in the network, there should be $(N-1)$ burst assembly queues and processors dealing with bursts towards the corresponding destinations.

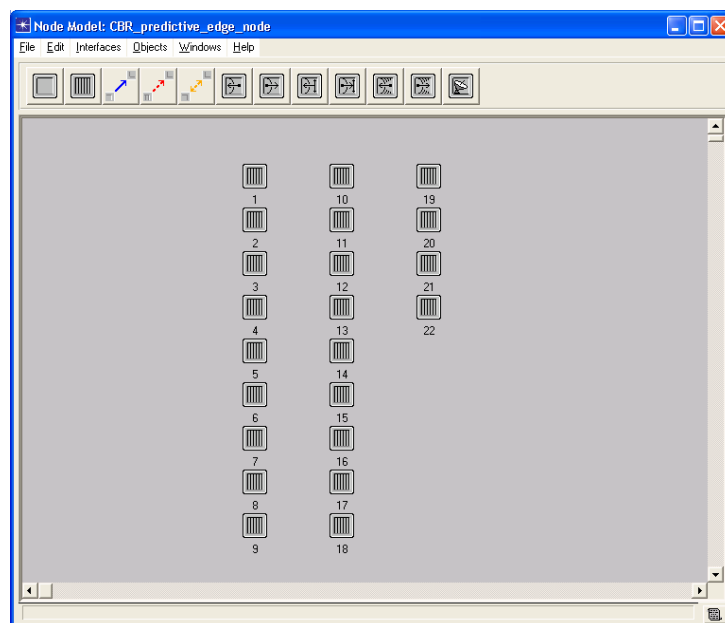


Fig 7-5: Node Model for edge node

The node model for each optical router is illustrated in Fig 7-5. As we can see, in each node

model there are 22 independent queue processes, representing the burst assembly queues and burst request processors for the designated destinations. These queue processes are intelligently activated by the program, as necessary. The structure for each of the queue process in the node model is the same, and is introduced in following sections.

7.1.3.1 Edge Node Process Model for WR-OBS and Traditional Reservation Schemes

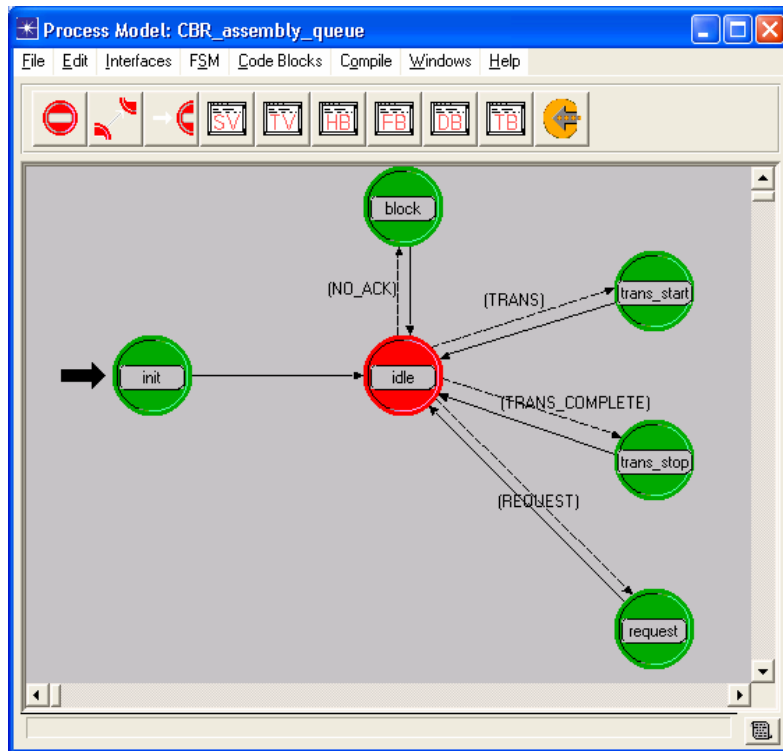


Fig 7-6: Edge node process model for WR-OBS and traditional reservation schemes

The queue process model inside the edge node model represents the burst assembly queue and burst processor for bursts towards a specific destination from the local edge node.

The structure of the queue process model is the same for both the WR-OBS and the traditional reservation schemes, as shown in Fig 7-6. The “init” state initializes the process model. For example, the “init” state starts the traffic generation as a background process, generating self-similar or Constant-Bit-Rate (CBR) traffic. The “init” state also schedules the first burst request after the simulation starts. The “idle” state acts as an intermediate state detecting the simulation events and transferring the simulation state into corresponding state.

The “request” state corresponds to the state where the lightpath request is sent from the edge node to the central node. In WR-OBS and traditional reservation schemes, the new burst request is sent after the corresponding burst is assembled for an initial duration T_{edge} , where the T_{edge} is determined by the capacity of burst assembly queue, the end-to-end delay criteria $T_{ete,max}$, the maximum delay allowed for the request in the central node $T_{sched,max}$, and the propagation time from the edge node to the central node T_{sig} . As we assume an infinite sized burst assembly queue, T_{edge} can be expressed as in Equation 7-5, where 0.0002 corresponds to 0.2 milliseconds as the maximum time required by the RWA processing.

$$T_{edge} = T_{ete,max} - 2T_{sig} - T_{sched,max} - 0.0002 \quad \text{Equation 7-5}$$

The “block” state is triggered by the arrival of a reject message from the central node. The reject message means that there are no lightpaths available for the current burst even if the burst is delayed till the end-to-end delay tolerance limit is reached. In this case, the whole content of the current burst is abandoned, and the burst is regarded as lost.

The process enters the “trans_start” state, when a positive lightpath acknowledgement is received by the edge node. As the lightpath has been reserved by the central node, the burst starts transmission when the OXC is configured after time T_{tune}^{50} .

Following the transmission of the current burst, the process enters the “trans_stop” state to release the lightpath resources and start a new burst aggregation process. Importantly, the difference between the WR-OBS and traditional reservation schemes mainly resides in this state.

More specifically, in WR-OBS, the time to stop the burst transmission is determined by the estimated burst length. Existing WR-OBS research assumes perfect knowledge of burst length or CBR traffic. Therefore, the “trans_stop” state schedules a lightpath release message to the central node after the completion of burst transmission at the first node and the burst propagation time, indicating the time to release the lightpath and update the global link-state

⁵⁰ It is assumed in WR-OBS that the propagation time from the central node to all the associated lightpath nodes is the same as the time from the central node to the edge node, such that the concerned OXC will be configured ready for the burst when the burst starts transmission.

database.

In traditional reservation scheme, the “trans_stop” state is triggered when the burst assembly queue becomes empty after the current burst starts transmission⁵¹. Then, a lightpath release message is scheduled after the burst aggregation time plus two times of the signalling propagation time from the edge node to the central node, such that the time required to propagate the lightpath release signal is included. Then the lightpath is eventually released and the central node global link-state database is updated at the scheduled time.

7.1.3.2 Edge Node Process Model for Pre-booking Scheme

The structure of the process model inside the pre-booking edge node is shown in Fig 7-7. The model comprises five states. The “init” state initializes the model, starts traffic generation, and schedules the first pre-booking request. The “idle” state captures the events and transfers to the corresponding event state.

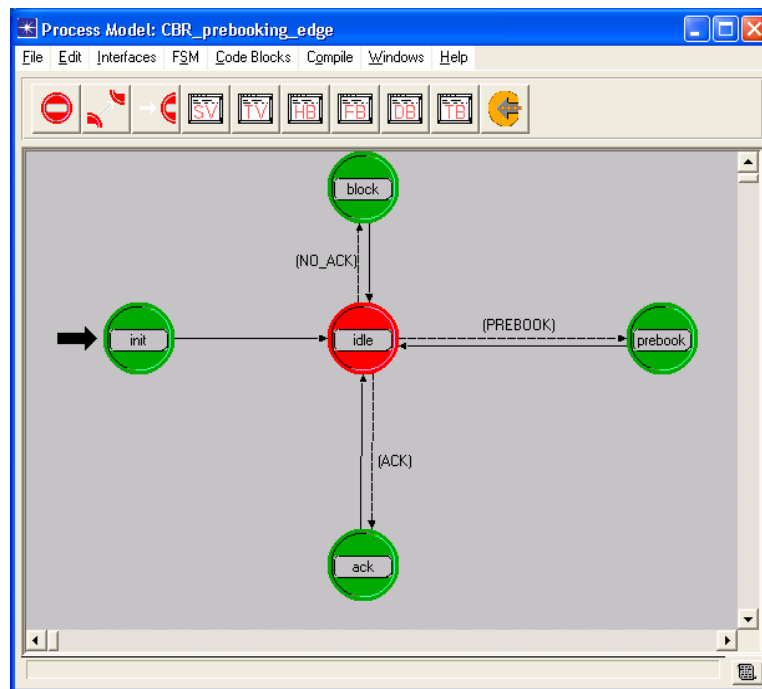


Fig 7-7: Edge node process model for pre-booking scheme

The “prebook” state is triggered by the start of the burst aggregation. In the pre-booking mechanism, the pre-booking is sent from the edge node to the central node as soon as the new

⁵¹ We use the UBS burst assembly mechanism in these simulation.

burst assembly starts, and the burst is handled by a “reserve-while-accumulating” principle. In the “prebook” state, traffic information such as the average incoming bit rate, the variance, and the delay information is put into the pre-booking request, and the request is sent from the edge node to the central node. The traffic statistics are obtained by observing the previous incoming traffic, and used by the central node to estimate the burst length⁵².

The “block” state handles the reject message from the central node. In this case, the burst is regarded as lost, the burst contents in the assembly queue are deleted, and a new burst aggregation starts when the next bit of traffic arrives.

The “ack” state responds to the lightpath acknowledgement from the central node. As the committed lightpath might be at a delayed time, the “ack” state advances the simulation time to the lightpath start holding time, and transmits the burst. The duration of the lightpath is determined by the central node based on the burst length prediction, so the lightpath will be torn down once the predicted burst transmission time expires. Therefore, if there are any residual bits in the burst assembly queue, those bits cannot be delivered and they will be abandoned due to the insufficient reservation. The analysis of bit loss rate due to insufficient reservations has been described in detail in Section 6.8.5. After the burst finishes transmission, the “ack” state starts a new burst aggregation and pre-booking cycle as soon as a new bit arrives.

7.1.4 Validation of the Simulation Models

All the models were validated before they were used to obtain the final simulation results. The self-similar traffic model was checked and evaluated via multiple mechanisms, including inspection of the generated traffic trace, the log-log time-variance plot.

The validation of the OPNET models was carried out via three approaches. The first approach was to debug the functions offline to ensure those functions acted correctly. The functions being debugged were the RWA algorithms used by the central node, the burst length estimation function, etc. The second approach was to embed functions into the simulation environment,

⁵² It is the central node who decides the start of burst transmission, so the burst aggregation time will not be known by the SP. Hence, the central node has to estimate the burst length based on the traffic statistics and the time when the burst aggregation starts as well as when the burst should start transmission.

and to monitor the behaviour of the functions in the simulation. The third approach was to trace the activities of the bursts and signalling messages, by running the simulation with a simple scenario with fewer events and just two network nodes.

7.2 Simulation Results for Self-Similar Traffic Traces

7.2.1 Simulated Burst Length Distribution

This section focuses on the burst length distribution, such as to compare the simulated results with the statistics provided in section 6.8.3.

The burst length distribution can be illustrated by the Probability Density Function (PDF) of the obtained traffic traces. In simulation, the PDF function is calculated from the Cumulative Distribution Function (CDF) of the obtained data according to Equation 7-6, where $y1$ and $y2$ are consecutive sample values of the burst length.

$$PDF(y1) = \frac{CDF(y2) - CDF(y1)}{y2 - y1} \quad \text{Equation 7-6}$$

In the simulation, traffic is generated from 10 and 100 ON-OFF independent sources, with each source arriving at 10Mbps incoming rate. The Pareto distribution parameters are the same for ON and OFF periods, where $\alpha = 1.5$ ($H = 0.75$), and $k = 0.0352$, which is equivalent to the time (in milliseconds) required to transmit a 44Byte packet with 10Mbps transmission rate on each stream [XVC00]. The total incoming bit rate is 50Mbps with 10 ON-OFF processes, and 500Mbps with 100 ON-OFF processes. The PDF for the aggregated traffic with aggregation time 10ms and 100ms are illustrated in Fig 7-8 (500000 aggregated bursts are generated in the simulation).

Fig 7-8 shows that when 10 ON-OFF processes are used, the burst length yields a bell-shaped distribution. However, the distribution does not strictly match the normal fit. Moreover, when the aggregation time increases to 100ms, the mismatch becomes apparent. When 100 ON-OFF processes are presented in the simulation, the resulted burst length distribution matches the normal distribution closely. At the backbone edge nodes, traffic is aggregated from a large number of users, so a large number of ON-OFF processes are required in simulation, such that

the behaviour of the high aggregation level achieved by having a large number of incoming users at the backbone edge can be truly simulated.

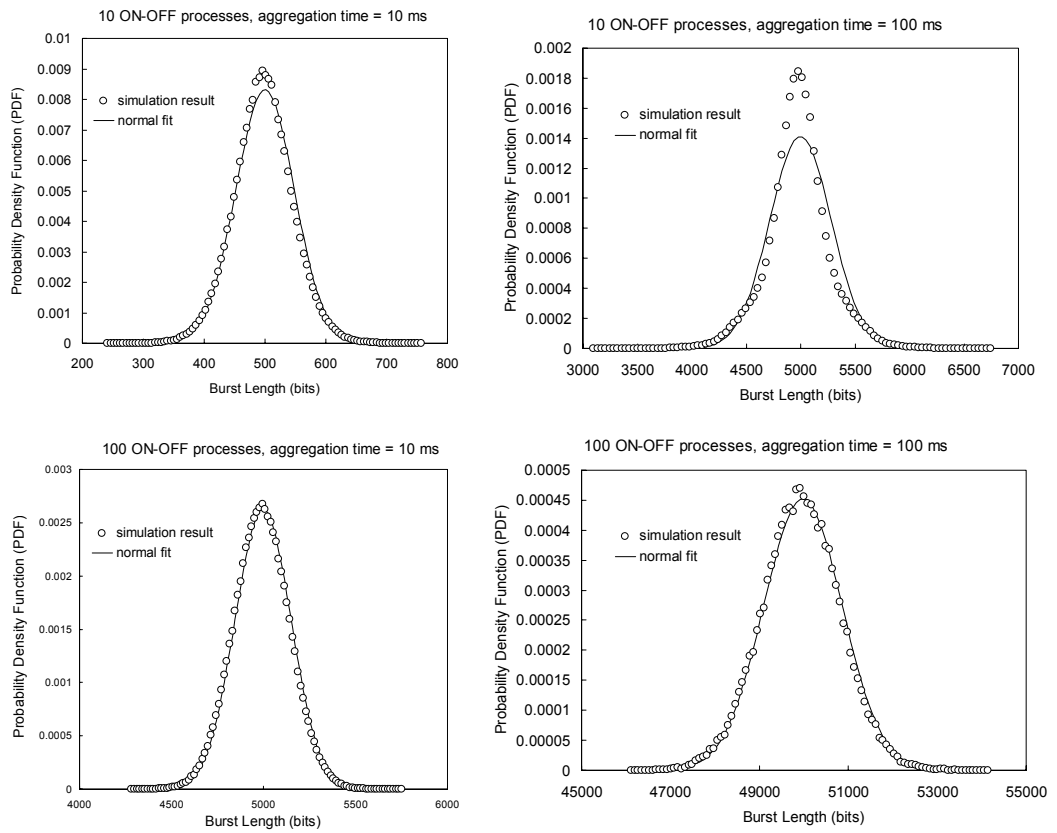


Fig 7-8: PDF for aggregation burst length from 10 and 100 ON-OFF processes

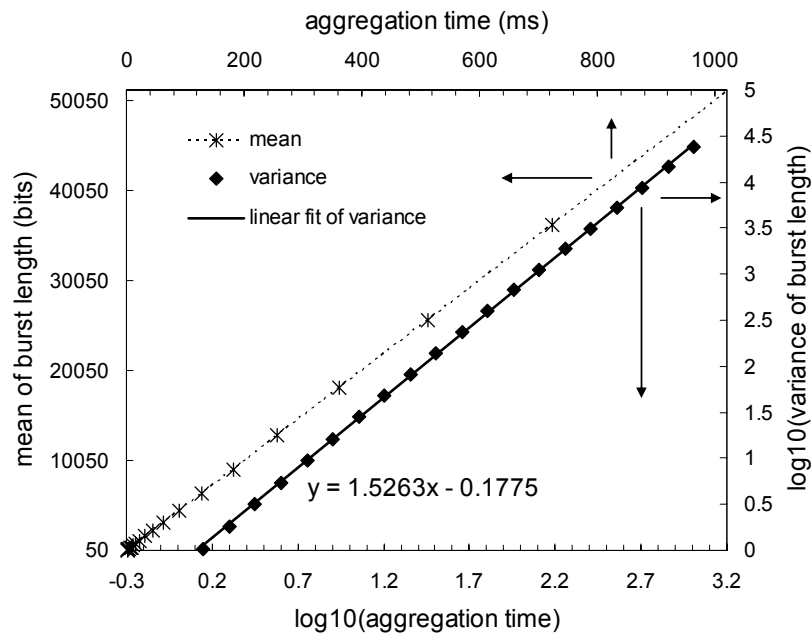


Fig 7-9: Mean and variance of aggregated burst length from 10 ON-OFF processes

Apart from the shape of the distribution, the mean value of the aggregated burst length should grow linearly with the aggregation time, and the logarithmic burst length variance should yield a linear relationship with the logarithmic aggregation time with a slope equivalent to $2H$ (According to Equation 6-16 and 6-17). This claim is validated by the simulation results, as shown in Fig 7-9.

7.2.2 Bit Loss Rate Caused by Insufficient Reservation

Simulations in this section aim to investigate the bit loss rate in relation to insufficient reservations caused by the prediction error. The theoretical bit loss rate based on the strict normal distribution is deduced and presented in Equation 6-25. The theoretical results are presented together with the simulation results, such as to cross-validate the results and to reflect the normal distribution factor of the simulated burst length.

Two groups of traffic traces are generated from 10 ON-OFF processes and 100 ON-OFF processes, respectively. Each ON-OFF source uses the same parameters for ON and OFF periods, and $\alpha = 1.5$, location parameter of the Pareto distribution $k = 0.0325$.

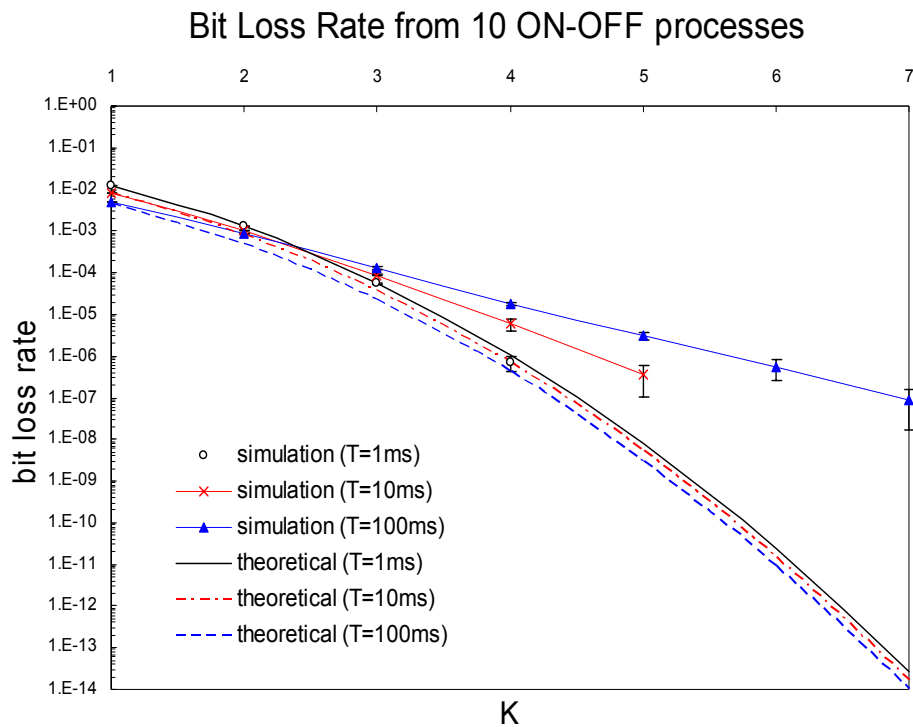


Fig 7-10: Bit loss rate from 10 ON-OFF processes

Fig 7-10 illustrates the bit loss rate from 10 ON-OFF processes. The horizontal axis is the k value in the proposed excess reservation strategy, as shown in Equation 6-18. The simulation results are presented with 95% confidence interval, and with the theoretical results. The results show that when the aggregation time is $1ms$, the simulation results are consistent with the theoretical results. However, when the aggregation time is $10ms$ or $100ms$ and K is larger than 2, the simulation results deviate from the theoretical results, being higher than the theoretical value. This can be explained by the imperfect match of the normal distribution to the aggregation burst length. Especially, when k value becomes larger than 3, the probability of a burst length being longer than the reserved duration becomes very small, so the bit loss rate becomes very sensitive to any mismatch between the simulation and the theoretical results. Additionally, as shown in Fig 7-8, the traffic generated from 10 ON-OFF processes does not strictly fit the normal distribution, so the gap between the simulation results and the analytical results becomes noticeable when k is large.

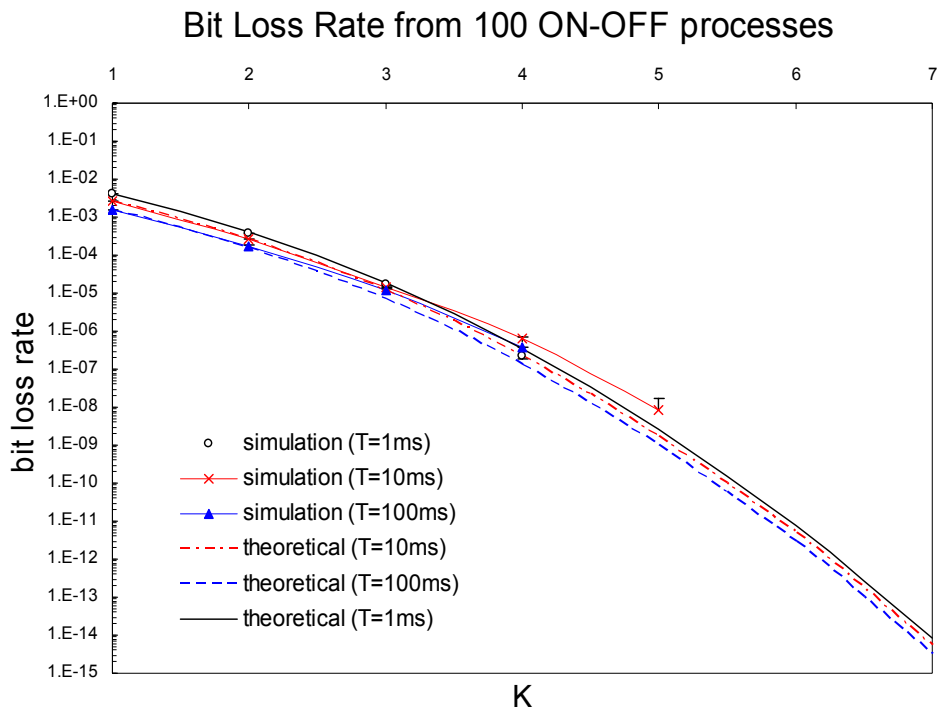


Fig 7-11: Bit loss rate from 100 ON-OFF processes

Nevertheless, the burst length distribution from 100 ON-OFF processes matches the normal distribution much better than the traffic from 10 ON-OFF processes, as shown in Fig 7-8. Therefore, it is expected that the bit loss rate of the traffic from 100 ON-OFF processes will

compare well with the theoretical results. Fig 7-11 presents the bit loss rate versus the k value. As anticipated, the simulation results match the theoretical results much better than the results from 10 ON-OFF processes, though there is still some deviation when k is larger than 3, which is mainly due to the high sensitivity when k is a large value.

As a conclusion, the bit loss rate decreases with the growth of the k value. However, due to the imperfect match in simulated burst length distribution, the simulated bit loss rate varies with the number of ON-OFF processes. A larger number of ON-OFF processes results in a closer match of the normal distribution, so the simulated bit loss rate will be closer to the theoretical results. Furthermore, the simulation results becomes very sensitive to the mismatch between the simulated and the theoretical distribution, when k is larger than 3.

7.3 Network Simulation Environment

Multiple scenarios have been simulated to investigate the performance of the three schemes in a national backbone environment. In all the scenarios, the USA NSFnet is used as the network topology, and it is presented in appendix B with the true fibre length in kilometres. Pennsylvania (PA) is selected as the location of the central node such that the maximum propagation delay from the edge nodes to the central node is minimized [MKB02].

The schemes are studied with both the CBR and the self-similar traffic. The same amount of traffic is generated for each source-destination pair. The traffic load is defined as the ratio of incoming traffic bit rate at each source-destination pair to the outgoing bit rate (wavelength transmission rate), where the outgoing bit rate is set to 10Gbps.

The time parameter associated with the OXC configuration and the laser tuning is set to 1 millisecond. The central node RWA processing time in the WR-OBS or conventional reservation schemes is a random value derived from a Beta distribution, bounded by 0 and 0.2ms, and with average 0.1ms. The RWA in the pre-booking scheme is set to three times of that in WR-OBS, and it is a random process following a Beta distribution, bounded by 0 and 0.6 ms, and with average 0.3ms. Additionally, the signalling and burst propagation time is calculated based on the factor that it takes 1 millisecond to propagate a signal for each 200 kilometres.

In terms of the burst aggregation, the UBS assembly mechanism is adopted in all the

simulations, because previous studies shows that UBS yields better blocking probability performance than the LBS approach [MDB01].

In all the simulations, random numbers with specified distributions are mainly produced by the random number generator in the OPNET simulation kernel, which is itself based on the BSD random number generator. The period of the BSD random generator is 2^{31} . Although this random generator does not produce high-quality randomness, it is acceptable if it is only used to introduce some variation instead of measuring the statistical quantities. Therefore, this random number generator is mainly used to activate the nodes at a random time, and to generate the stochastic value for RWA execution time. With regard to the self-similar traffic generation, a high-quality random number generator is required. As the Mersenne Twister pseudo-random number generator has a period length of $2^{19937}-1$, and it provides very good randomness [MT98], it is adopted in our simulation to generate self-similar traffic.

All the results are obtained from the average of 10 independent experiments, and the results are presented showing the 95% confidence interval.

7.4 Network Simulation Results with CBR Traffic

With CBR traffic, the burst length can be accurately calculated based on the incoming bit rate and the burst aggregation time. Therefore, no prediction error is involved in the simulation. The results mainly demonstrate the effects of efficient resource utilization in terms of average blocking probability and lightpath bandwidth efficiency.

Two major scenarios have been studied in the simulations. The first scenario illustrates the blocking probability where each burst is only given one opportunity to find a lightpath at the specified time. If the lightpath is available at the requested time, the burst will be delivered to the other end. Otherwise, if no lightpath is free at the requested time, the burst will be blocked. The second scenario aims to show the capability of supporting latency-sensitive traffic. Each burst has a strict end-to-end delay requirement, and the lightpath provisioning can only be delayed to a limited degree, such as to increase the chance of getting a free lightpath whilst experiencing extra delay within the tolerable range.

7.4.1 Results for Scenario-1 with a Single RWA Attempt

For the first scenario, where only single RWA attempt is offered by the central node, the time for initial burst aggregation $T_{edge} = 0$ in WR-OBS and traditional reservation schemes. The reason for this is to keep the burst length in these two schemes almost the same as that in the pre-booking mechanism. As there is no specific deadline for the request, the request will be buffered and scheduled in the central node based on the first-come-first-served principle.

More specifically, in WR-OBS and traditional reservation schemes, the request will only be considered once while they are processed by the RWA processor. If no lightpaths are available, the request will be rejected. In simulations with the pre-booking mechanism, the RWA algorithm will only search for a lightpath at the earliest burst transmission time, $t_{earliest}$, where $t_{earliest}$ is the time when the lightpath acknowledgement is received by the edge node, and it is expressed in Equation 7-7. t_{start} is the time that the burst starts aggregation and the pre-booking request is sent to the central node. T_{sig} is the signalling propagation time from the edge node to the central node, or reversely. T_{RWA} is the time spent on the RWA algorithm.

$$t_{earliest} = t_{start} + 2T_{sig} + T_{RWA} \quad \text{Equation 7-7}$$

The three schemes are investigated from various perspectives, showing the burst blocking probability versus varying loads, number of wavelengths, and network size, respectively.

7.4.1.1 Burst Blocking Probability versus the Traffic Load

This set of experiments demonstrates the burst blocking probability performance in relation to the traffic load. The simulation assumes the original size of the NSFnet, and 12 wavelengths on each link. The traffic load increases from 0.1 to 0.9.

The results are shown in Fig 7-12. It reveals that the performance of the traditional reservation mechanism and WR-OBS are much worse than that of the pre-booking mechanism. The traditional reservation mechanism yields the highest burst blocking probability. The WR-OBS is slightly better than the traditional reservation scheme, and the pre-booking mechanism significantly outperforms the other two. For instance, when the load is 0.1, the pre-booking approach yields little loss at about 10^{-6} , whilst that of the other two schemes is as high as 10^{-1} . When the traffic load increases, the blocking probability for all three schemes increase. Due to

the high traffic demand when the load approaches 0.9, the network becomes very congested, where resources are nearly saturated, hence, the performance of all three schemes are similar.

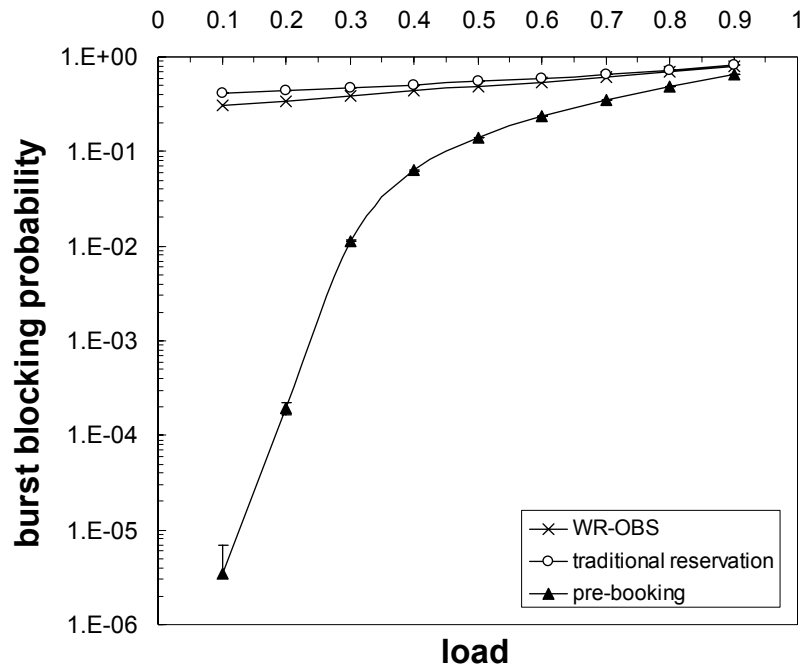


Fig 7-12: Burst Blocking Probability versus Traffic Load for Scenario-1

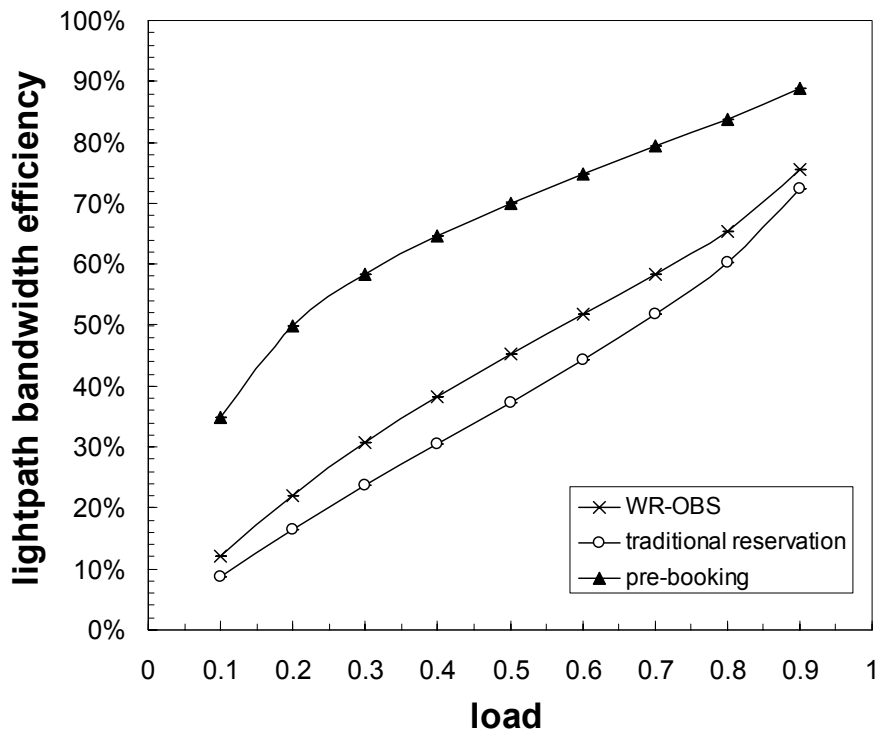


Fig 7-13: Lightpath Bandwidth Efficiency versus Traffic Load for Scenario-1

The corresponding lightpath bandwidth efficiency is shown in Fig 7-13. As anticipated, traditional reservation yields the lowest lightpath bandwidth efficiency, because it idly holds lightpath resources during the signalling propagation time in both the lightpath acknowledgement and the lightpath release stages. WR-OBS is more efficient, because it only idly holds resources in the lightpath acknowledgement stage. Pre-booking performs the best, because it does not idly hold any resources during the signalling propagation. Meanwhile, the lightpath bandwidth efficiency improves with increasing traffic load, which means that the efficiency improves for larger burst lengths.

Fig 7-12 and Fig 7-13 indicate that the network blocking probability performance is inevitably linked with the efficiency of network resource utilization. The more efficient the network resources are utilised, the more traffic that can be supported, and the lower is the blocking probability.

7.4.1.2 Burst Blocking Probability versus the Number of Wavelengths

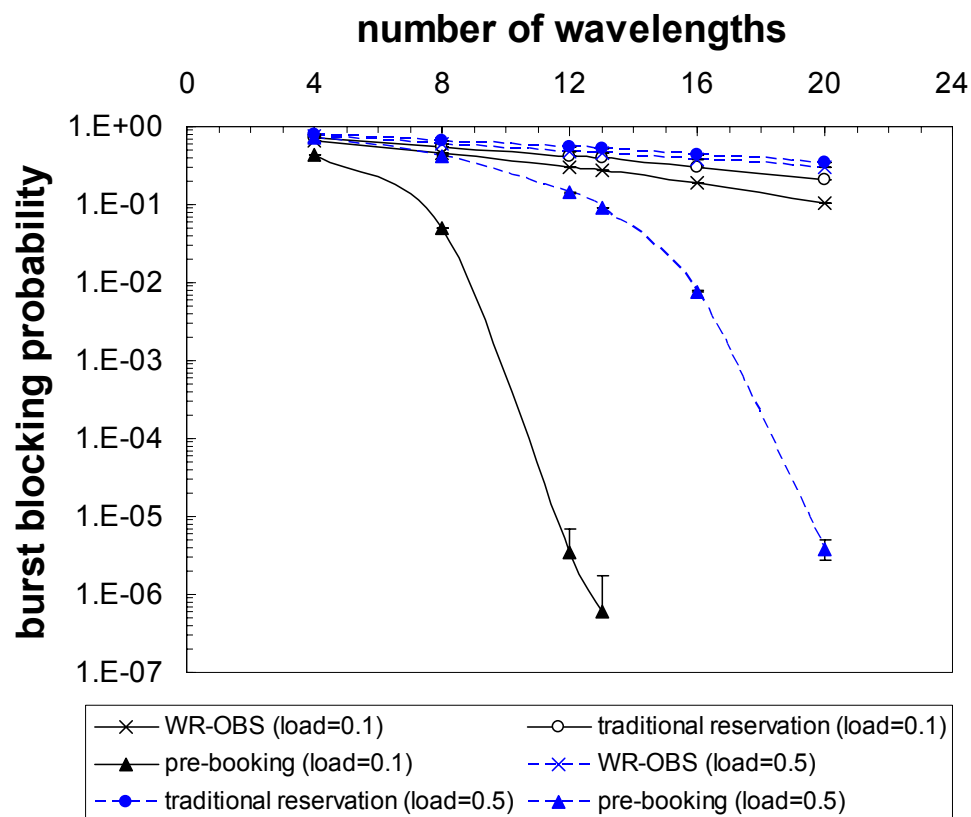


Fig 7-14: Burst Blocking Probability versus the Number of Wavelengths for Scenario-1

In this set of experiments, the simulation is still based on the original NSFnet, but the number of wavelengths on each link is increased from 4 to 20. The results, as shown in Fig 7-14, are obtained for traffic loads of 0.1 and 0.5, respectively.

The results show that the burst blocking probability decreases when more wavelengths are provided. The burst blocking probability for the pre-booking mechanism decreases at a much faster rate than that for the other two mechanisms. Particularly, when 13 wavelengths are offered, the pre-booking mechanism achieves a burst blocking probability smaller than 10^{-6} at a load of 0.1⁵³.

The conclusions are similar to those drawn from Fig 7-12, and the pre-booking mechanism can utilise network resources in the most efficient way such as to achieve the lowest burst blocking probability.

7.4.1.3 Burst Blocking Probability versus the Network Size

In this set of experiments, 12 wavelengths are provided on each link, but the network size is scaled by a factor such as to investigate the relationship between the burst blocking probability and the network size.

Fig 7-15 shows the results for traffic loads at 0.1 and 0.5, respectively. The network size on the horizontal axis is the scaling factor between the simulation network size and the original network size. For example, 0.25 means that the network size used in simulation is actually 0.25 times the original NSFnet size. The results demonstrate that the burst blocking probability for WR-OBS and the traditional reservation scheme dramatically increases with the growth of network size. This is mainly due to the increased signalling and burst propagation time in an enlarged network, and the lightpath is idly held longer in a larger network. The pre-booking mechanism achieves the best performance. For instance, when traffic load is 0.1, zero blocking probability is observed for pre-booking approach when network sizes smaller than the original

⁵³ A previous study shows that 13 wavelengths are just enough to simultaneously support one connection for each source-destination pair in NSFnet. In our simulation, due to the un-timely lightpath release caused by the signalling mechanism, a second lightpath might be required to be setup while the existing connection is still being held. Therefore, each source-destination pair may demand multiple connections at the same time. Thus, 13 wavelengths are not sufficient to support all the connection requests, causing a non-zero blocking probability.

NSFnet size. However, the corresponding blocking probability for the other two schemes are much higher than that of the pre-booking scheme.

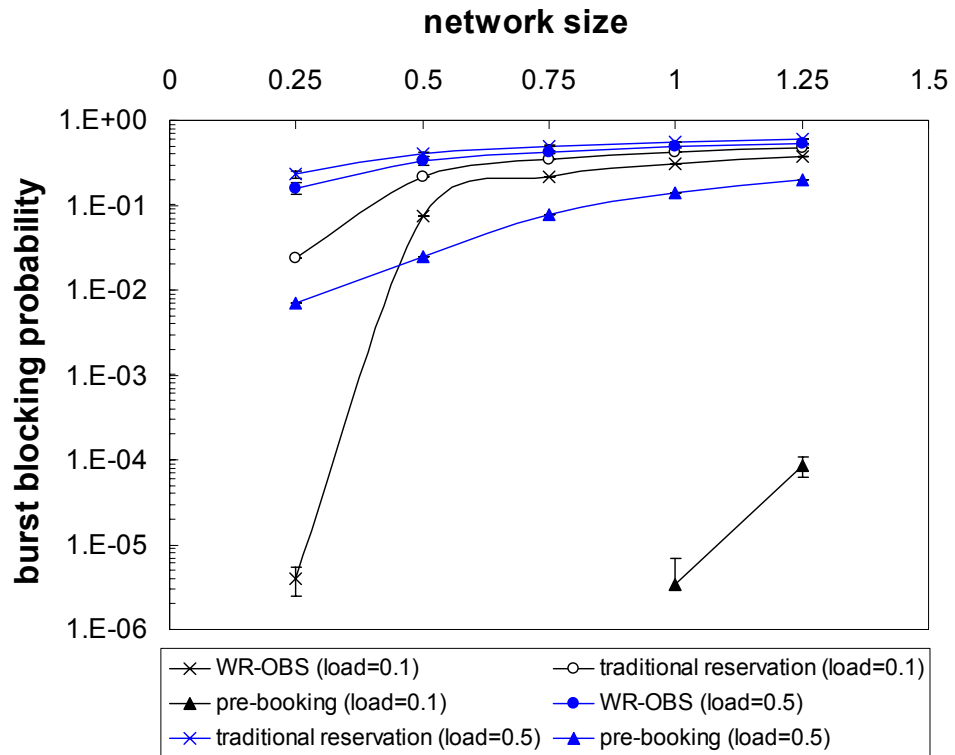


Fig 7-15: Burst Blocking Probability versus the Network Size for Scenario-1

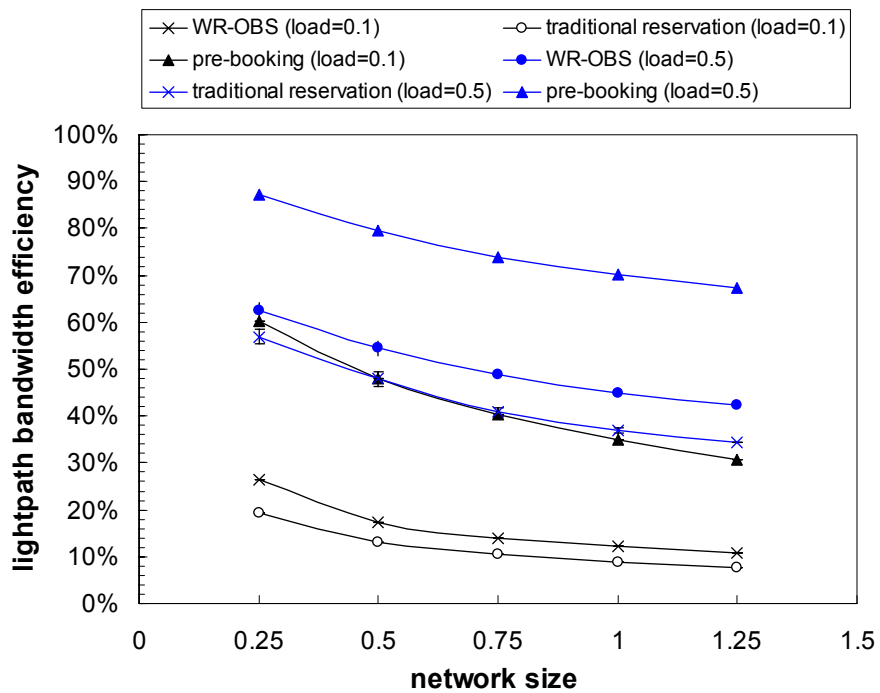


Fig 7-16: Lightpath Bandwidth Efficiency versus the Network Size for Scenario-1

The lightpath bandwidth efficiency is shown in Fig 7-16. As expected, the lightpath bandwidth efficiency decreases when the network size gets larger, being consistent with the analysis, and also provides an explanation for the burst blocking probability shown in Fig 7-15.

The network size impacts the signalling propagation delay in a straight line. In WR-OBS and traditional reservation mechanisms, the longer is the signalling propagation delay, the longer the network resources will be unnecessarily held. Therefore, network size influences the lightpath bandwidth efficiency and the blocking probability.

7.4.2 Results for Scenario-2 with Multiple RWA Attempts

In real network applications, each request is usually given multiple opportunities to attempt to find a free lightpath with the appropriate delay. Additionally, as more and more forms latency-sensitive multimedia traffic are being transported via the Internet, it is becoming very important to consider the network performance in terms of support for latency-sensitive traffic with specified delay tolerances.

The simulations for scenario-2 place strict constraints on the end-to-end delay and the blocking probability. The maximum end-to-end delay was set to be 90 milliseconds, which is a typical requirement for videoconference services [MKB02]. The target maximum average burst blocking probability was set to 10^{-4} to meet current ITU recommendations.

7.4.2.1 Time Parameters for End-to-End Delay

The end-to-end delay is the time that elapses from the arrival of first bit of the new burst at the edge burst aggregation queue, until the arrival of the first bit at the destination node. The total end-to-end delay for each burst in WR-OBS and the traditional reservation schemes can be decomposed into several parts as shown in Equation 7-8.

$$T_{ete} = T_{edge} + T_{sig} + t_{sched} + t_{RWA} + T_{sig} + t_{tuning} + T_{ete-prop} \quad \text{Equation 7-8}$$

Where T_{edge} is the initial burst aggregation time before the lightpath request is sent, T_{sig} is the signalling propagation time from the source to the central control node, or in reverse direction, t_{sched} is the total time that the request spends in the central scheduling queue, t_{RWA} is the time for RWA processing, t_{tuning} is the time for OXC laser configuration (set to be 1millisecond), and

$T_{ete-prop}$ is the end-to-end propagation time for each burst.

To ensure that the burst is delivered within the specified delay tolerance, WR-OBS and the traditional reservation schemes can mainly rely on the control of T_{edge} , $t_{sched, max}$. where $t_{sched, max}$ is the maximum central node scheduling time, determining whether the request should be rejected if the request has stayed in the central node for too long and the end-to-end delay criteria is violated. Let $t_{RWA, max}$ be the maximum possible time for RWA processing, and $T_{ete-prop, min}$ be the minimum end-to-end propagation delay for the specified source-destination pair. The value for T_{edge} and $t_{sched, max}$ must satisfy Equation 7-9 such that the end-to-end delay is guaranteed within the tolerance $T_{ete, max}$.

$$T_{ete, max} \geq T_{edge} + T_{sig} + t_{sched, max} + t_{RWA, max} + T_{sig} + t_{tuning} + T_{ete-prop, min} \quad \text{Equation 7-9}$$

Similarly, in the pre-booking mechanism, as the initial burst aggregation time is always zero, $T_{edge} = 0$, the end-to-end delay is mainly controlled by the earliest burst transmission time, $t_{earliest}$, and the parameter, called max_delay , which is used by the central node to determine the maximum edge delay allowed for the request such that the end-to-end delay criteria is satisfied. The corresponding time parameters for pre-booking approach should satisfy Equation 7-10.

$$T_{ete, max} \geq (t_{earliest} - t_{start}) + max_delay + t_{tuning} + T_{ete-prop, min} \quad \text{Equation 7-10}$$

Where $t_{earliest}$ is the earliest burst transmission time, and t_{start} is the time when the first bit of the new burst arrives at the assembly queue. As the burst transmission can only start after the lightpath acknowledgement is received, $t_{earliest}$ should satisfy Equation 7-11.

$$t_{earliest} \geq t_{start} + T_{sig} + t_{RWA, max} + T_{sig} \quad \text{Equation 7-11}$$

Furthermore, in order to achieve fairness among the requests for different source-destination pairs, the central node scheduling principle based on the EDF is studied in [MKB02]. More specifically, the fairness issue arises because the propagation delay is dependent on the location of the source, destination and central nodes. Taking WR-OBS as an example, if the source node is close to the central node, the signalling delay, T_{sig} , will be shorter than that from a more remote source. Given that all the requests yield the same end-to-end delay tolerance $T_{ete, max}$, if T_{edge} is set to the same for all the source-destination pairs, the maximum time allowed for central

schedule, $T_{sched, max}$, varies with T_{sig} and $T_{ete-prop, min}$. However, a longer $T_{sched, max}$ indicates that the request can be delayed for longer, thus increasing the chance of successfully getting a free lightpath, which eventually causes unfairness among the requests for different source-destination pairs.

Therefore, setting the same value for $T_{sched, max}$ (or max_delay in pre-booking scheme) belonging to various source-destination pair will provide equal opportunities for all the requests in terms of getting a free lightpath. The simulation in this thesis fixes $T_{sched, max}$ for all the source-destination pairs, whilst the value for T_{edge} (or $t_{earliest}$ in pre-booking approach) is adjusted based on T_{sig} and $T_{ete-prop, min}$ such that Equation 7-10 and 7-11 are fulfilled. (More details of the scheduling principle can be found in [MKB02].)

7.4.2.2 Burst Blocking Probability versus the Load

This set of simulations mainly demonstrates the burst blocking probability versus the varying traffic load. The original NSFnet size is assumed in the simulation, and 12 wavelengths are provided on each link.

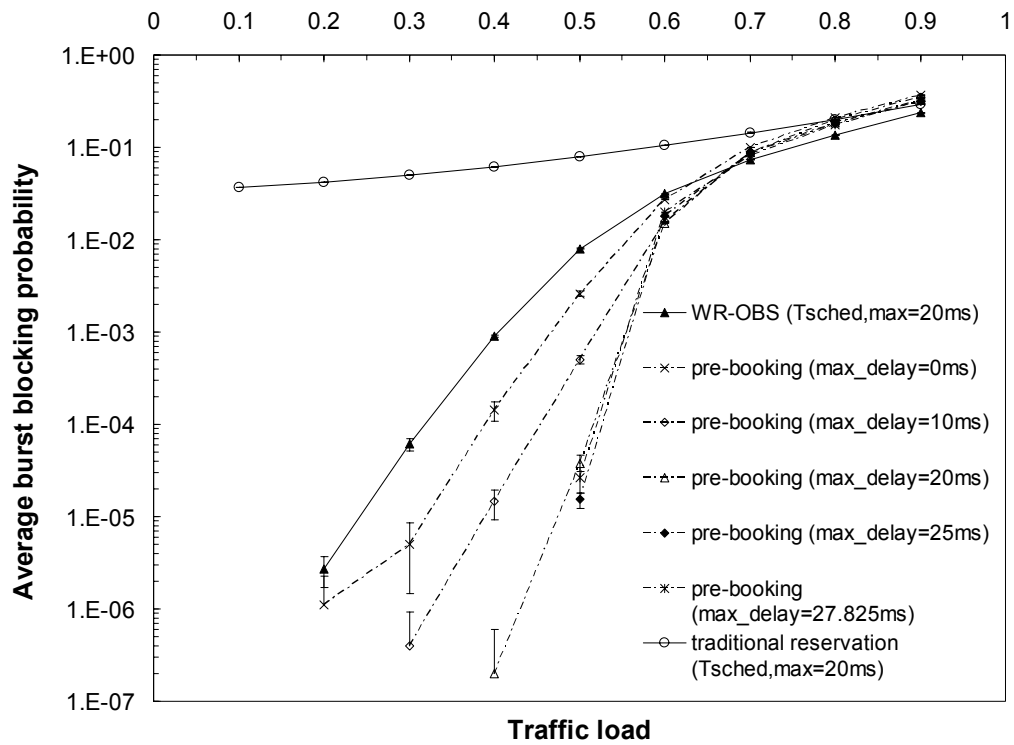


Fig 7-17: Burst Blocking Probability versus the Load for Scenario-2

Fig 7-17 shows the average burst blocking probability for the three schemes. The results for WR-OBS and the traditional reservation schemes are the optimal case with the maximum central node scheduling time $T_{sched,max}=20ms$.

The results clearly illustrate that with our proposed pre-booking mechanism, the best performance is obtained when max_delay is 25ms, and a traffic load of nearly 0.55 can be transmitted with the targeted blocking probability of 10^{-4} . Whilst, with WR-OBS, the maximum supported traffic load reduces to about 0.35. The traditional reservation scheme yields the worst performance, and an extremely high blocking probability. Among the three schemes, pre-booking produces the most economical reservations, and provides the best support for the traffic with specific blocking probability and end-to-end delay requirements.

Meanwhile, if the max_delay parameter increases in pre-booking scheme, the time allowed for reattempting a lightpath increases, so the blocking probability reduces. However, the lowest blocking probability is not achieved with the highest max_delay value, similarly to the results published in ref [MKB02]. This is due to a tradeoff between the burst aggregation time and the lightpath delay time max_delay . If the burst aggregation time increases, the lightpath is more efficiently utilized, so the blocking probability reduces. However, with fixed end-to-end delay tolerance in our simulation, the increase of max_delay results in the decrease of burst aggregation time. Therefore, there is an optimum value of max_delay to tradeoff the burst aggregation time and the max_delay .

Furthermore, the proposed blocking probability for pre-booking mechanism increases very quickly above a load of 0.55, and yields a marginally higher blocking rate than the WR-OBS at a high traffic loads beyond 0.9. This is due to the effects of the RWA algorithm used in the pre-booking mechanism. As the RWA algorithm reserves the lightpath in future time slots, the remaining wavelength resources become fragmented. When the traffic load becomes exceptionally high and the burst length increases, these available wavelength resource fragments are too short to be useful, and cannot be productively used for lightpaths.

With regard to the average lightpath bandwidth efficiency, the results are presented in Fig 7-18. The results for WR-OBS and the traditional reservation scheme are the optimal case with the maximum central node scheduling time $T_{sched,max}=20ms$.

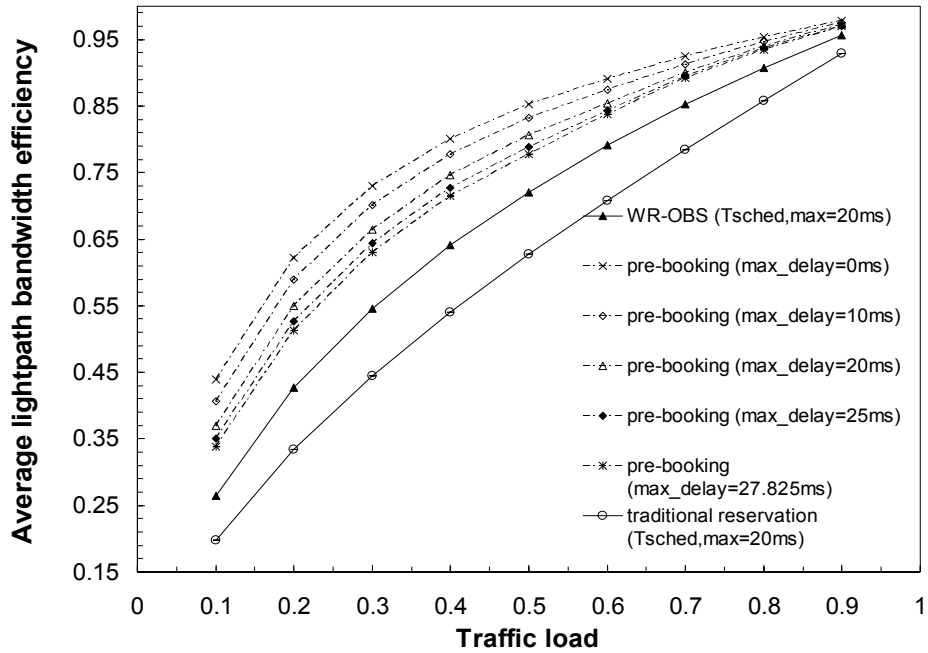


Fig 7-18: Lightpath Bandwidth Efficiency versus the Load for Scenario-2

As anticipated, the traditional reservation approach yields the lowest average lightpath bandwidth efficiency, and pre-booking achieves the best performance. Fig 7-18 also illustrates that the efficiency improves with increasing traffic load, which means that the efficiency grows for larger burst lengths. In the pre-booking mechanism, a longer max_delay results in lower bandwidth efficiency. This is because the increased max_delay causes decreased initial burst aggregation time, the burst length decreases, and the lightpath bandwidth efficiency decreases as a consequence.

7.4.2.3 Burst Blocking Probability versus the Number of Wavelengths

Fig 7-19 shows the burst blocking probability with the increasing number of wavelengths in the original NSFnet network size. $T_{sched,max}=20ms$ for WR-OBS and traditional reservation schemes, and $max_delay = 25ms$ for the pre-booking scheme, and traffic load is 0.1 and 0.5, respectively.

Results demonstrate that when the wavelength number is as low as 4, there is little difference between the three schemes, because the network resources are insufficient, and the saving of resources from the signalling propagation time can hardly overcome the absence of network resources. When more wavelengths are available, the performance gain of pre-booking becomes apparent and significant. When the traffic load is 0.1 and 8 wavelengths are provided, the

pre-booking mechanism yields a burst blocking probability of about 10^{-5} , whilst, that for WR-OBS and traditional reservation is about 10^{-2} and 10^{-1} , respectively. Similarly, when the load is 0.5, pre-booking mechanism achieves almost zero-blocking probability with 13 or more wavelengths. However, for WR-OBS approach, more than 16 wavelengths are required to achieve the same performance. For the traditional reservation approach, over 20 wavelengths have to be provided to achieve non-blocking performance.

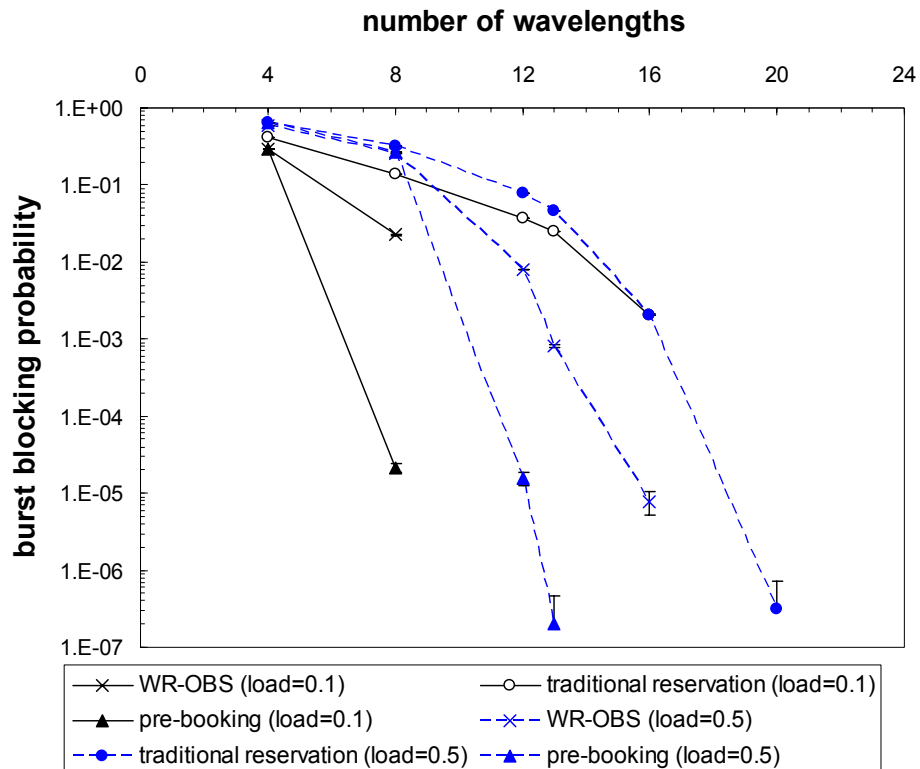


Fig 7-19: Burst Blocking Probability versus the Number of Wavelengths for Scenario-2

7.4.2.4 Burst Blocking Probability versus the Network Size

This set of simulation investigates the effects of network size on the burst blocking probability for bursts with strict end-to-end delay requirements. The number of wavelengths is set to 12 for all the links in the NSFnet, and the network size is represented by the scaling factor, which is ratio of simulation network size to the original NSFnet size.

Additionally, as the network size is associated with the propagation time, the value for $t_{sched, max}$ or max_delay should be adjusted with the network size. For the original NSFnet size, each

source-destination pair yields a total minimum signalling and end-to-end propagation time, $T_{prop, sd}$, as shown in Equation 7-12, and the maximum $T_{prop, sd}$ among all the source-destination pair is denoted as $Max(T_{prop, sd})$.

$$T_{prop, sd} = 2T_{sig} + T_{ete-prop, min} \quad \text{Equation 7-12}$$

Thus, for WR-OBS and traditional reservation schemes, $t_{sched, max}$ is decided based on Equation 7-13, where $scale$ is the scaling factor for the network size. Similarly, for the pre-booking scheme, max_delay is calculated in the same way.

$$t_{sched, max} = T_{ete, max} - scale \times Max(T_{prop, sd}) - t_{RWA, max} - t_{tuning} \quad \text{Equation 7-13}$$

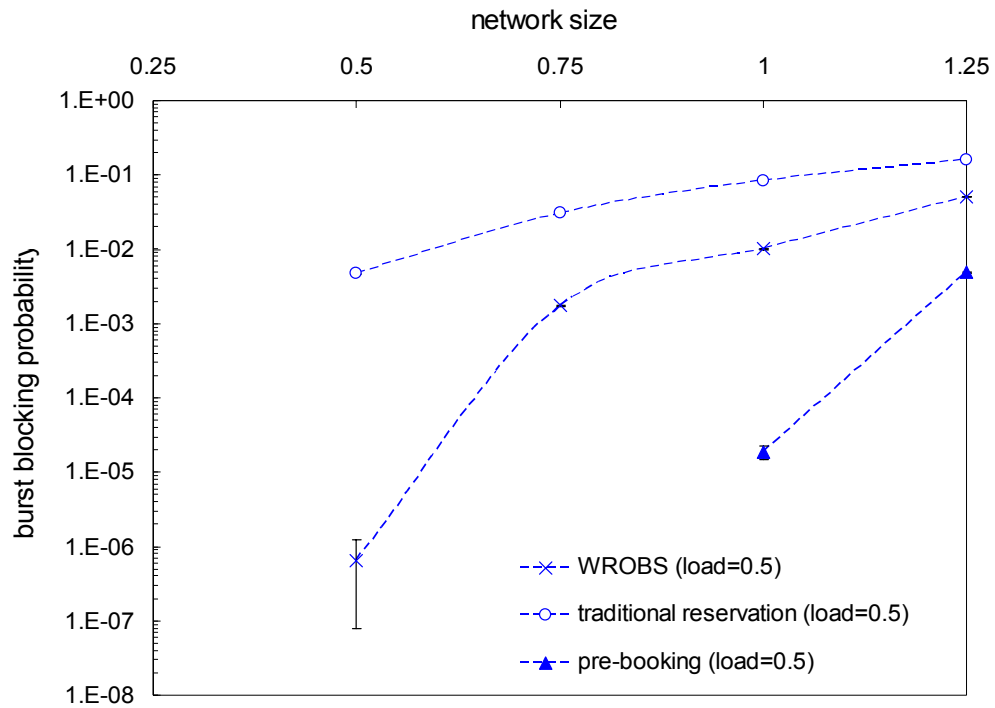


Fig 7-20: Burst Blocking Probability versus the Network Size for Scenario-2

Fig 7-20 illustrates the burst blocking probability versus the network size, when traffic load is 0.5.

The results show that the burst blocking probability increases with the network size. When load is 0.5, the blocking probability for traditional reservation approach is higher than 10^{-3} for network sizes larger than 0.5 times of the original NSFnet. WR-OBS yields lower blocking rate, but its blocking rate reaches about 10^{-3} for a network size at 0.75 times of the original NSFnet.

The pre-booking approach achieves the best performance, and network size can be approximately extended to 1.25 times of the original network size for a blocking probability of 10^{-3} .

7.5 Network Simulation Results with Self-Similar Traffic

With self-similar traffic, the duration of lightpath holding time relies on the burst length prediction in our proposed pre-booking mechanism. Traffic loss is not only caused by the absence of lightpaths, but is also caused by prediction errors. Assuming the lightpath has been established via the pre-booking procedure, if the reservation duration is shorter than the actual burst length, the residual bits in the burst cannot be successfully transmitted to the other end; otherwise, if the reserved duration is longer than the actual burst length, the whole burst is shipped to the destination, but some of the resources will be unnecessarily held, reducing the lightpath bandwidth efficiency.

For WR-OBS, existing studies assumes perfect knowledge of the burst length, and the simulation in this thesis keeps this assumption, such that the pre-booking mechanism can be compared with the best performance of the WR-OBS mechanism. With regard to the traditional reservation scheme, all the OXC operations depend on the receipt of signalling messages, so there is no prediction error involved.

In the simulations, each self-similar traffic source is produced from 100 independent ON-OFF processes. All the ON-OFF processes use the same parameters for ON and OFF periods, where $\alpha = 1.5$ ($H = 0.75$), and the minimum ON or OFF duration is equivalent to the time for transmitting 4.4Kbytes⁵⁴ traffic on each stream. The traffic load is adjusted by the transmission rate for each ON-OFF stream, based on Equation 7-4.

Simulations with self-similar traffic takes considerable time, so results are mainly collected for

⁵⁴ In current data networks, 50% of packets are 44 bytes long [XVC00]. However, using 44 bytes minimum ON-OFF period requires extraordinary long time to simulate a short simulation. For example, it takes about 45 minutes to finish a 5-second long simulation with a 1Gbps arrival rate. Setting 4.4Kbytes minimum ON or OFF period accelerates the simulation to a speed of about 25 minutes for a 200-second long simulation with a 1Gbps arrival rate. Furthermore, in [DB02], the simulation results are obtained by assuming a minimum 5Kbytes ON or OFF period, so our simulation setting is similar to this publication.

scenario-2 as with the CBR results, where multiple RWA attempts are permitted for each request such that a lightpath is provided with appropriate delay. All the parameters are kept the same as the simulations with CBR traffic, i.e. $T_{sched, max} = 25ms$ for WR-OBS and traditional reservation schemes, and $max_delay = 20ms$ for the pre-booking scheme.

Additionally, as the pre-booking mechanism is affected by the prediction error, the simulation results not only reveal the burst blocking probability and the lightpath bandwidth efficiency for the pre-booking mechanism, but also present the bit blocking probability arising from bit loss caused by the prediction error. All the performance measures are investigated using various traffic loads and varying k parameter values as described in the excessive reservation strategy.

7.5.1 Performances versus the Traffic Load

This set of results mainly illustrates the network performance with varying traffic load. Each burst has a strict 90 milliseconds end-to-end delay requirement, so the results demonstrate the network capability of supporting latency-sensitive traffic in the original-sized NSFnet with 12 wavelengths on each link.

7.5.1.1 Burst Blocking Probability

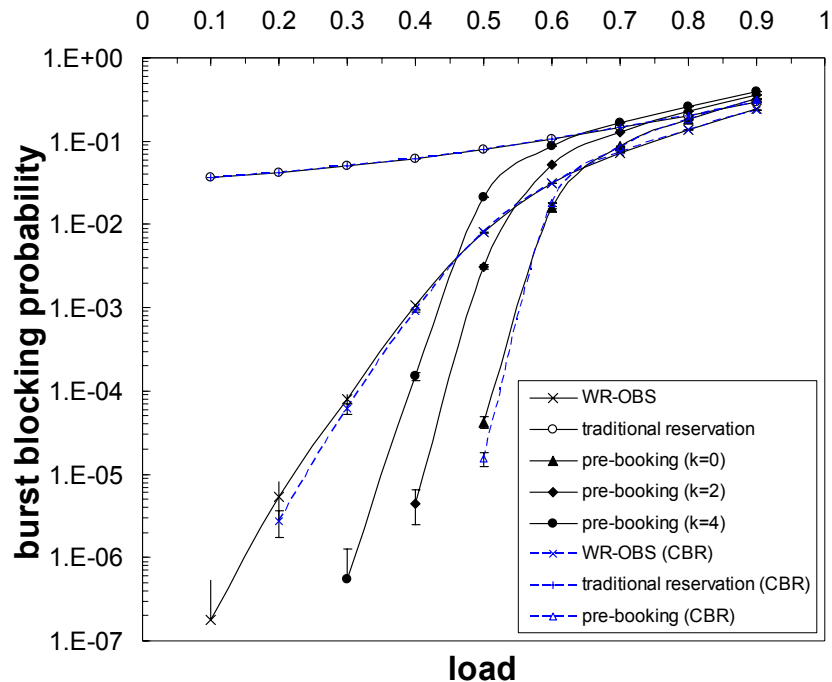


Fig 7-21: Burst Blocking Probability versus the Traffic Load with Self-similar

Traffic

Fig 7-21 shows the burst blocking probability performance with self-similar traffic. Additionally, in order to reveal the influence of self-similarity on the blocking probability, the corresponding results with CBR traffic are also presented with blue dashed lines and points.

The results clearly show that for WR-OBS and traditional reservation schemes, the burst blocking probability with self-similar traffic is very close to that obtained with CBR traffic. For pre-booking mechanism, the burst blocking probability with $k = 0$ is close to the corresponding blocking rate with CBR traffic. This indicates that the influence of self-similarity is negligible on the blocking probability, and this is because the pre-booking takes the average burst length in the reservation, and the actual requested load is the same as the CBR traffic case.

However, when k increases, the burst blocking probability increases for the pre-booking mechanism. This is because when $k = 0$, the average reserved load is equivalent to the offered load. When k increases, the average reserved load becomes higher than the offered load, hence, the burst blocking probability increases accordingly.

7.5.1.2 Bit Blocking Probability

However, does the burst blocking probability reflect the true traffic loss for the pre-booking mechanism? The answer is No. Specifically, in the pre-booking mechanism, the traffic loss is affected by both the lightpath availability and the prediction error. Though a lightpath is provided by the pre-booking procedure and the burst is regarded as successfully delivered, if the reservation duration based on the burst length prediction is shorter than the actual burst duration, residual bits may still be lost due to insufficient reservations. Therefore, the bit loss rate for pre-booking mechanism is a combination of the burst blocking and insufficient reservation durations caused by the prediction error.

Fig 7-22 supplies detailed bit blocking probability results for the three schemes. Basically, four groups of results are presented. The first group is the overall bit loss rate for all the three schemes with self-similar traffic. The second group is the corresponding bit loss rate with CBR traffic, which is mainly presented by the blue dashed lines and points. The third group is the bit loss rate caused by the lightpath blocking for the pre-booking mechanism with self-similar

traffic, as shown in green dashed lines and points. The fourth group is the bit loss rate caused by insufficient reservations for the pre-booking mechanism with self-similar traffic, as shown in red dashed lines and points.

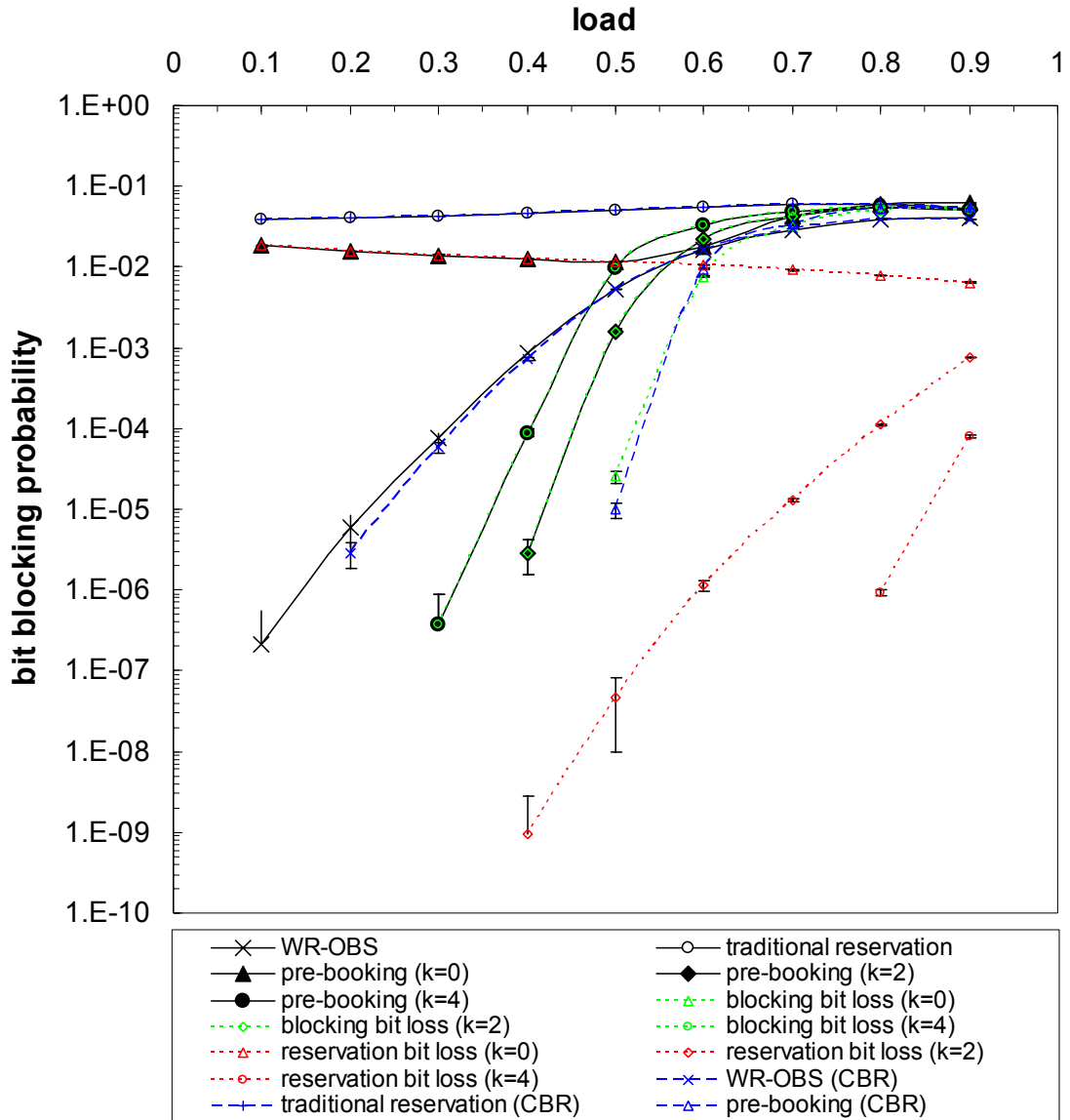


Fig 7-22: Bit Blocking Probability versus the Traffic Load with Self-similar Traffic

First of all, the overall bit loss rate for the WR-OBS and the traditional reservation schemes with self-similar traffic is close to the results with CBR traffic, indicating that the self-similarity has a limited influence on bit blocking probability performance.

However, for the pre-booking mechanism with self-similar traffic, the bit loss rate is greatly affected by the k value. When $k = 0$, the bit loss rate for pre-booking is a little bit unusual, as it

is when the traffic load less than 0.5. Surprisingly, the bit loss rate remains at a level as high as 10^{-2} . In fact, this result is reasonable. As mentioned before, the bit loss rate for the pre-booking mechanism is a combination of lightpath blocking and insufficient reservation durations. The overall bit blocking probability for pre-booking can be decomposed to the “blocking bit loss” and the “reservation bit loss”, as shown in the corresponding green and red lines and points. When $k = 0$ and the load is less than 0.5, the blocking bit loss rate is nil, whilst, the reservation bit loss rate is very high. Therefore, the overall bit loss rate is dominated by the reservation bit loss rate. When the load grows beyond 0.5, the bit loss caused by the lightpath blocking increases dramatically, so the overall bit loss rate is then dominated by the blocking bit loss rate. Similarly, when k increases to 2 and 4, the corresponding reservation bit loss rates are very small, and the overall blocking bit loss rate is mainly caused by the lightpath blocking.

As anticipated, the bit loss rate caused by insufficient reservations decreases with increasing k value. When $k = 0$, only the averaged burst duration is reserved for each burst, whilst, about 50% of bursts will yield a burst length longer than the average length. Therefore, the reservation bit loss remains high when $k = 0$. When k increases, the probability that the actual burst length is greater than the reserved length decreases, so the reservation bit loss rate drops. Additionally, when $k = 0$, the reservation bit loss rate decreases slowly with the increasing traffic load. This is consistent with the analytical results as shown in Fig 7-11 and Fig 7-10, where the bit loss rate caused by prediction error decreases slowly when the burst length or burst aggregation time increases. However, when $k = 2$ or $k = 4$, the simulated reservation bit loss rate increases with the growth of traffic load. This can be explained by the inaccurate match with the Gaussian distribution for bursts aggregated from the ON-OFF superposition model. As shown in Fig 7-11 and Fig 7-10, the simulated bit loss rate cause by insufficient reservations deviates from the theoretical results when k increases.

More importantly, for the pre-booking mechanism, the ability to support traffic loads with a specified blocking probability does not simply increase with increasing k value. For example, when $k = 0$, due to the high reservation bit blocking rate, the bit loss rate is very high even when the traffic load is very low, so the mechanism can hardly support a traffic load of 0.1 with bit loss rate less than 10^{-4} . When $k = 2$, the reservation bit loss rate decreases dramatically, and the overall bit loss rate is dominated by the lightpath blocking. The scheme in this case shows the

ability to support a traffic load of up to 0.45 with overall bit loss rate less than 10^{-4} . However, when $k = 4$, the performance decreases again. Though the bit loss rate caused by insufficient reservation keeps decreasing to a very low level, the bit loss caused by the lightpath blocking increases quickly, and the supported traffic load decreases to about 0.4 with the targeted overall bit loss rate. The result indicates that there is an optimal value for the k parameter such as to maximize the supported traffic load with a specified bit loss rate. When k is too small, the bit loss rate caused by insufficient reservations is high, resulting in high overall bit blocking probability. When k is too large, the reserved load becomes significantly higher than the actual offered load, and the bit loss rate caused by the lightpath blocking becomes dominant, thus causing a high overall bit blocking probability as well.

7.5.1.3 Lightpath Bandwidth Efficiency

In terms of lightpath bandwidth efficiency, the results are demonstrated in Fig 7-23.

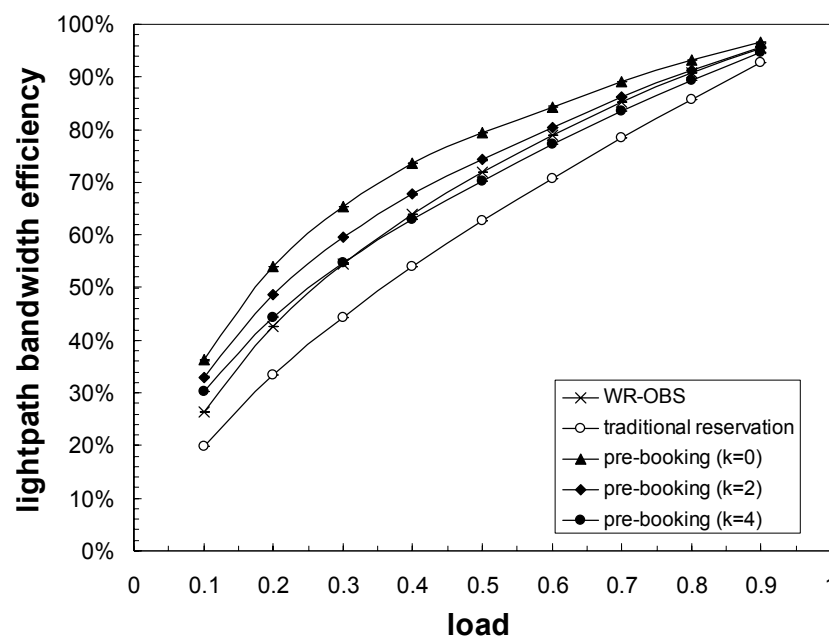


Fig 7-23: Lightpath Bandwidth Efficiency versus the Traffic Load with Self-similar Traffic

Similarly, due to the increased burst length, the lightpath bandwidth efficiency is enhanced by the increased traffic load. The pre-booking mechanism yields a higher lightpath bandwidth efficiency than WR-OBS and the traditional reservation schemes when k is no more than 2. Moreover, the traditional reservation scheme possesses significantly lower lightpath bandwidth

efficiency than the other two schemes. The reason is mainly due to the reservation mechanism adopted. The less resources are unnecessarily held, the higher the lightpath bandwidth efficiency will be.

For the pre-booking mechanism, the lightpath bandwidth efficiency decreases with increased k value. This is reasonable because the larger the k value is, the longer the reservation duration will be, and the higher will be the probability that the actual burst duration is shorter than the reservation duration. Therefore, the lightpath bandwidth efficiency is reduced due to excessive reservation strategy adopted. Particularly, when $k = 4$, the lightpath bandwidth efficiency becomes even lower than WR-OBS when the load is high, degrading other performance measures such as the burst and bit blocking probability. Hence, it is important to choose an appropriate k value, such that the excessive reservation strategy does not go beyond the network capabilities and does not degrade the overall bit blocking probability.

7.5.2 Performance versus the k Value for Pre-booking Mechanism

As the k parameter in the proposed excessive reservation strategy plays an important role in determining the network performance, this section concentrates on showing the performance with various k values.

7.5.2.1 Burst Blocking Probability

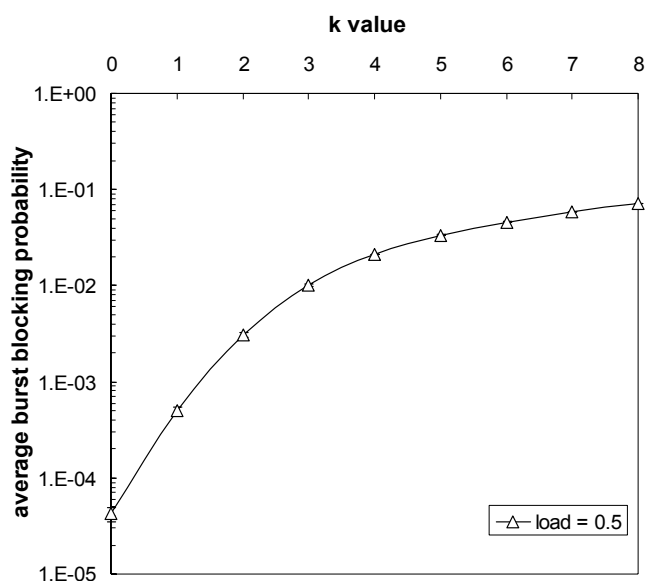


Fig 7-24: Burst Blocking Probability versus the k Value in Pre-booking

Mechanism with Self-Similar Traffic

Fig 7-24 shows the burst blocking probability versus the k value, when the traffic load is fixed at 0.5. The affect of the k value on the burst blocking probability is clearly illustrated. The burst blocking probability increases when the k value increases. This is because that when k grows, the reserved traffic loads increase, causing more burst blocking.

7.5.2.2 Bit Blocking Probability

Fig 7-25 demonstrates the overall bit loss rate for traffic load 0.5, with k varying from 0 to 8. Each overall bit loss rate is decomposed into the corresponding blocking bit loss rate and the reservation bit loss rate, such as to better appreciate the reasons for the overall bit loss.

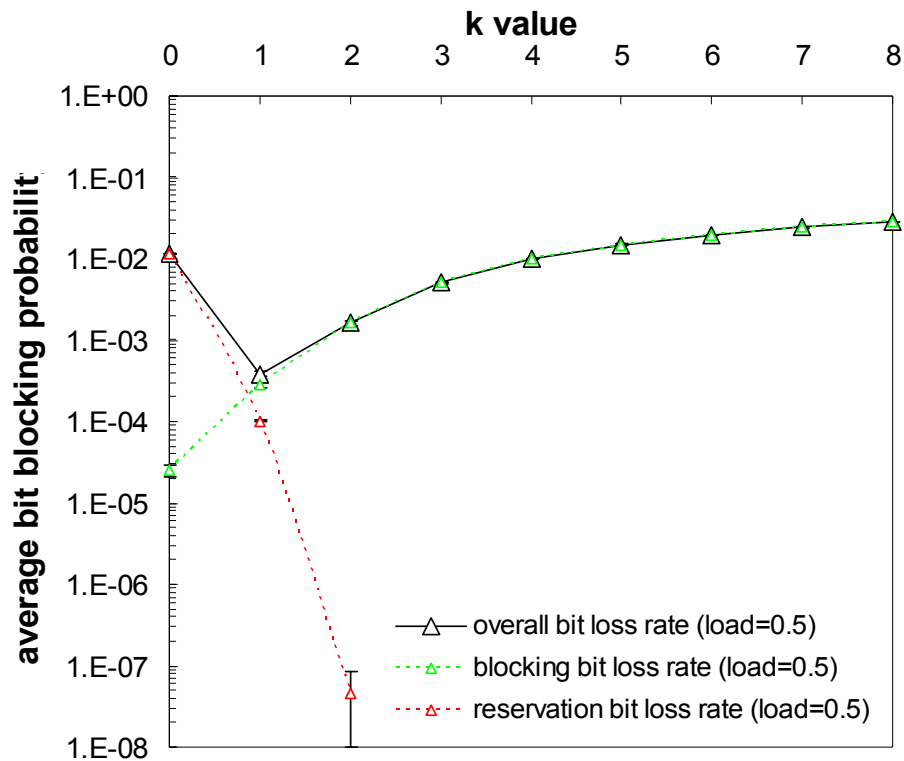


Fig 7-25: Bit Blocking Probability versus the k Value in Pre-booking Mechanism with Self-Similar Traffic

As it shows, the overall bit loss rate decreases first with k no larger than 1, and then it climbs up gradually with the increasing k. This can be explained by the corresponding blocking bit loss rate and the reservation bit loss rate. As demonstrated, the reservation bit loss rate is about 10^{-2} when k is 0; it decreases to approximately 10^{-4} when k is 1, then it further reduces to nearly 10^{-7}

or nil when k is 2 or more. The decreasing reservation bit loss rate indicates that the bit loss caused by insufficient reservation reduces if the k value increases. However, a larger k value can cause increased lightpath blocking. Therefore, increasing k value can cause contradictory affects on the reservation bit loss and blocking bit loss. There is a best k value where the lowest overall bit loss rate is achieved by best balancing the reservation bit loss and the blocking bit loss.

7.6 Summary

This chapter studies the simulation models and results for pre-booking, WR-OBS, and the traditional reservation mechanisms in connectivity acknowledged OBS networks. This chapter also provides a detailed description for self-similar traffic models.

The bit loss rate caused by insufficient reservation in the pre-booking mechanism is simulated and discussed assuming the burst length follows normal distribution. Both the simulation and theoretical results show that the bit loss rate decreases quickly with the increase of the k value, validating the analysis in section 6.10.

The network simulation results for the pre-booking mechanism were obtained for both the CBR and the self-similar traffic in NSFnet. With CBR traffic, two scenarios were developed. The first scenario shows the performance when each burst is only given one chance of RWA calculation. The second scenario shows the performance when supporting latency sensitive traffic. The overall conclusion for CBR traffic is that the pre-booking mechanism yields the lowest blocking probability and can support the highest load for latency-sensitive traffic. The burst blocking probability increases with increasing traffic load, whilst, it decreases when more wavelengths are provided on each link. As the network size affects the propagation time, the larger is the network size, the higher will be the burst blocking probability, the lower will be the ability to support latency-sensitive traffic with specified burst blocking probability, and the lower will be the lightpath bandwidth efficiency.

For self-similar traffic, the results only demonstrate the ability to support latency-sensitive traffic. The performance for the pre-booking mechanism is affected by the k parameter as employed in the excessive reservation strategy. Specifically, the burst blocking probability increases with the increasing k , because the reserved load increases with the k parameter.

The bit blocking probability includes the traffic loss caused by both the lightpath blocking and the insufficient reservation, and it provides a better measurement on traffic loss. The results show that the bit loss rate caused by insufficient reservation decreases with increasing k . However, the overall bit loss rate does not simply decrease with the growing k value. When k is small, the reservation bit loss is high, and so the overall bit loss rate is high. When k is large, the lightpath blocking becomes severe, and the overall bit loss rate reaches a high level as well. Therefore, there is an optimal value for the k parameter such as to maintain the reservation bit loss at a low level, whilst not severely affecting the lightpath blocking.

In addition, the k parameter affects the lightpath bandwidth efficiency in the pre-booking mechanism. The larger is the k parameter, the more resources will be unnecessary held, thus degrading the lightpath bandwidth efficiency. Therefore, it is very important to control the k parameter in the pre-booking mechanism when prediction error is involved.

There are other methods to predict the burst length, but all the methods incur prediction errors. The excessive reservation mechanism provides an approach to compensate for the effects of the prediction error, and a careful selection of k value can help to achieve the lowest traffic loss.

Chapter 8 Conclusions

Current commercial core optical networks employ semi-permanent lightpaths (static-WRON) to provide end-to-end connections between locations. Due to their static nature, it is easy and cheap to operate such a network. However, in the future, network traffic volume is likely to be much higher than current levels, and multimedia services will become an important element. These changes are likely to have an impact on the future core optical network, where the traffic will become bursty and the bandwidth demand will be high. Static WRON does not adapt to the traffic fluctuations, and can lead to poor resource utilization.

The situation can be improved by introducing dynamic provisioning and statistical multiplexing into the network operations. Though the dynamic lightpath establishment is becoming a realistic proposition via automated management systems, in a connection-based network bandwidth-on-demand services may cause long delays to other applications when some of the sessions are long and the network is busy. One of the possible solutions is to reduce the transmission granularity by packetalizing the traffic and so allow for interleaving. OPS, OBS, and WR-OBS are typical propositions that possess features of packetalizing, statistical multiplexing, and dynamic lightpath provisioning; they mainly differ in the transmission granularity and the corresponding signalling procedures.

OPS is the optical version of traditional packet switching. It is the architecture with the finest transmission granularity, providing a high degree of flexibility and efficiency in terms of resource utilization. However, OPS is impeded by slow optical device developments, and is unlikely to be commercially deployed in the near future.

By increasing the transmission granularity, OBS and WR-OBS both realize the concept of OPS in a more realistic manner. With current technology, it is feasible to deploy OBS and WR-OBS in many areas, where each end user is offered with a high bandwidth in Gbit/sec region, and multimedia and real-time services are easily accessible. However, OBS and WR-OBS still contain some drawbacks. For instance, OBS imposes certain hardware challenges such as the need for flexible wavelength converters. Additionally, OBS suffers from high blocking probabilities, and requires further research to satisfactorily support it. WR-OBS relies on a

centralized control plane, which is vulnerable to scalability issues. Furthermore, it does not support latency-sensitive traffic well due to inefficiencies in the reservation mechanism.

Proactive resource reservation forecasts future traffic events, so creating opportunities for better network planning and network optimization. The work presented in this thesis considers improvements for a variety of OBS architectures and critically evaluates the benefits and limitations of each approach.

Based on the initial idea of pro-actively reserving network resources for anticipated future traffic, this research proposes a corresponding pre-booking mechanism for the classical OBS networks, a hybrid network architecture with both the WR and OBS operating, and connectivity acknowledged OBS networks, respectively. Combined with the pre-booking mechanism, this research also studies the RWA problem to achieve reduced blocking probabilities for a given traffic matrix, as well as providing various traffic statistics to help with predictive reservation.

8.1 Discussion

8.1.1 Pre-booking Mechanism for Classical OBS Networks

The pre-booking mechanism in classical OBS networks encourages customers to book network resources before the actual event takes place. At the same time, intelligence in each SP also anticipates prospective traffic demands based on historical records.

The proposition employs various means of intelligence, including traffic prediction and negotiation with the CNO, to pre-book resources. To improve the efficiency of the resource allocation process, intelligence, such as the path estimation function based on the link cost, is introduced at the CNO edge nodes.

Preliminary results show that the simple pre-booking mechanism is able to reduce the blocking probability for pre-booked traffic. However, this causes biased resource allocation for pre-booked traffic, and leaves non-predicted traffic little resource.

The situation can be improved in several ways. For example, the SPs can try to negotiate with the CNO to see if by delaying the traffic for a short time network resources will become available. Alternatively, the CNO can update the routing table based on the existing

commitments, such that the traffic is sent via various less utilized routes.

8.1.2 Pre-booking Mechanism for a Hybrid Network Architecture and RWA Mechanism

Pre-booking is applied in a different way in the hybrid network architecture. The SPs in this architecture do not predict any specific bursts. Instead, they frequently forecast the average traffic load for the various source-destination pairs, and obtain a number of end-to-end lightpaths from the CNO. The establishment of pre-booked lightpaths inherits the spirit of wavelength routing. At the same time, it offers a classical OBS service for delivering transient traffic with “best effort” support.

Simulation results show that the proposed network architecture achieves a lower blocking probability than for both the pure classical OBS networks and pure static WRON architectures, assuming no wavelength conversion is supported. However, if wavelength conversion support is provided for the OBS services, the new approach does not provide advantages over the pure classical OBS network scheme. The study confirms that classical OBS greatly relies on the wavelength conversion technology to avoid high blocking probabilities. If wavelength conversion is not available, classical OBS generally performs poorly.

The second major topic addressed is the RWA algorithm for the wavelength constrained Max-RWA problem. The problem is important in order to achieve the minimum blocking probability for a traffic matrix associated with wavelength constrained WDM networks. This research proposes a two-stage RWA approach to produce an optimal or near-optimal solution in a short time frame. Numerical results illustrate that the proposed RWA mechanism has for the first time solved popularly cited traffic matrices with proven optimality in a few milliseconds or seconds. Future work is required to refine the mechanism such that it is suitable for various network topologies and matrices.

8.1.3 Pre-booking Mechanism for Connectivity Acknowledged OBS Networks

Traditional connectivity acknowledged OBS networks employ a two-way acknowledgement protocol to setup a lightpath for each burst. The optical switches are configured when they receive the corresponding signalling messages during the acknowledgement stage. As the

signalling requires time to propagate from one node to another, some resources are actually idly held during the signalling propagation time. In large backbones, the signalling propagation time is long, resulting in more resources being unnecessarily held, which eventually limits latency-sensitive service provisioning.

The pre-booking mechanism is able to pro-actively schedule the switching action at the time when burst is being aggregated. By excluding the signalling propagation time in lightpath holding time, much more resources become available for the payload transmission in large backbones, thus improving the overall network performance.

The pre-booking mechanism requires a technology to predict the burst length. With widely observed self-similar traffic, this work follows the common view that the aggregated burst length is normally distributed. Accordingly, the work proposes an excess reservation strategy, where the reserved duration is the mean burst duration plus k times its standard deviation. Use of the k factor in the reservation strategy reduces the probability that the real burst length is larger than the reserved length.

The simulation results for CBR traffic show that the pre-booking mechanism yields the lowest blocking probability, the highest lightpath bandwidth efficiency, and the highest supported traffic load for latency-sensitive services, among the three schemes under study.

For self-similar traffic, the performance is affected by the k parameter associated with the excessive reservation strategy. In other words, the performance of the pre-booking mechanism can be degraded by an inappropriate selection of the k value, where too many or too little resources are reserved. More specifically, if the k parameter is too small, the bit loss caused by insufficient reservation is high, thus making the total loss high. If the k parameter is too large, the reservation load becomes high, and the burst blocking rate increases, which also causes a high overall loss. Therefore, there is an optimal k value that is a trade-off between the reservation loss caused by the insufficient pre-booking and the lightpath loss caused by high reservation loads.

8.2 Future Work

This work can be extended in the following areas:

- ◆ Traffic prediction plays an important role in the pre-booking mechanism. It is necessary to study the traffic characteristics over both small time scales and large time scales, such that an appropriate prediction methodology is used for the corresponding application. The pre-booking mechanism needs to be further investigated with various prediction approaches imbedded.
- ◆ The pre-booking mechanism in classical OBS networks requires further development in terms of edge intelligence. For example, the path estimation function for the pre-booking mechanism has yet to be implemented. The function could make use of the link cost gradient to calculate paths for each pre-booking request, where the link cost gradient in the edge link-state database owned by the CNO edge node can be intelligently updated based on the existing commitments, current pre-booking requests, and future trends. Additionally, the offsetting the burst launch time using a *shift* scheme can be further studied to discover the extra edge delay cost introduced by the shifting process.
- ◆ For classical OBS networks and hybrid network architectures, OBS routing can be dynamically updated based on the existing reservation commitments and the congestion status. A routing table based on the shortest paths is simple, but is inflexible. Future work should introduce co-operation between the pre-booking mechanism and the OBS routing table maintenance, such that the OBS flows are intelligently delivered via various routes.
- ◆ For the hybrid network architecture, future work can further investigate the effects of various pre-booking intervals for specific traffic profiles. Additionally, more simulations could be performed to study the wavelength requirements for hybrid network architectures. The current results were obtained with 32 wavelengths on each link. However, there might be a point, beyond which the proposed architecture performs better than classical OBS networks for the case when wavelength conversion is supported.
- ◆ The two-stage RWA approach can be further studied and refined such as to accommodate more instances of network topologies and traffic matrices. Additionally, subject to the performance of the algorithm, an improved RWA can be further adapted to accommodate dynamic lightpath provisioning in MPLS/GMPLS networks.
- ◆ In the hybrid network architecture, various RWA approaches can be applied and compared.

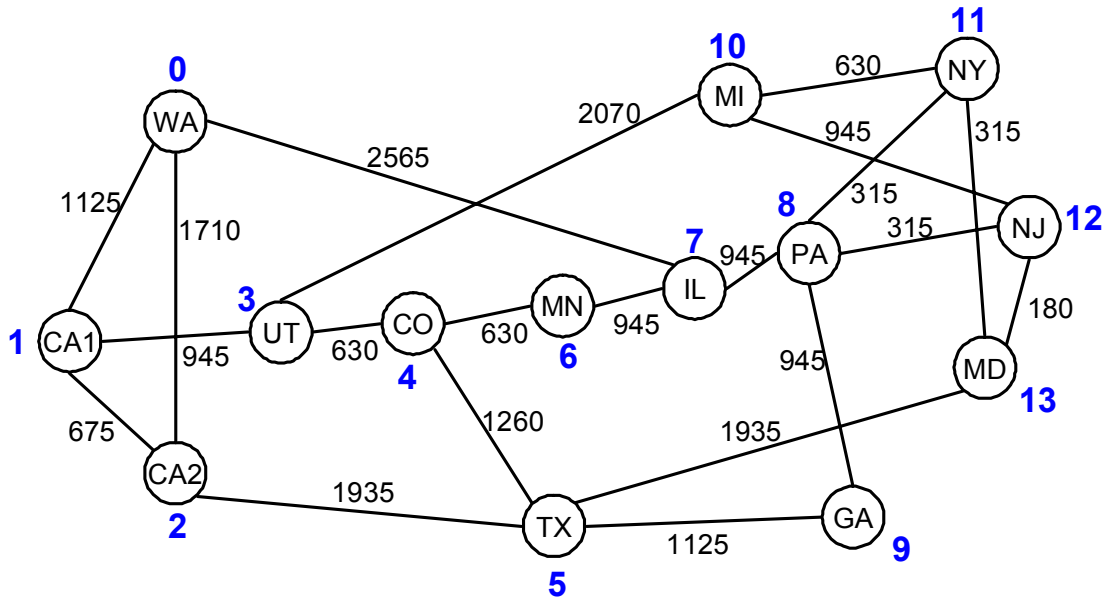
The aim of this future work is to reveal the impact of different RWA algorithms on measures of network performance such as the blocking probability and the wavelength requirements.

Appendix A – Author’s Publications

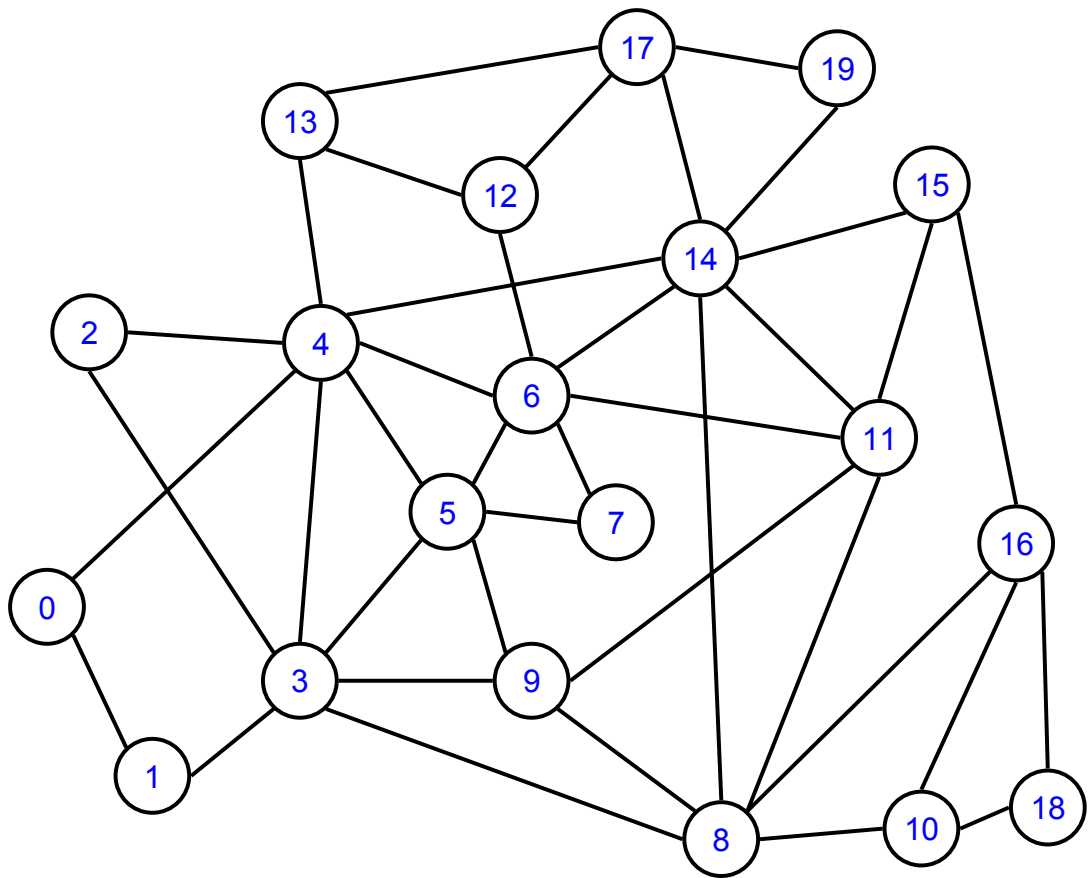
- [1] Huifang Kong and Chris Phillips, “Pre-booking Reservation Mechanism for Next Generation Optical Networks”, accepted for publication in 2006 Optical Communications Issue of the Journal of Selected Topics in Quantum Electronics, IEEE, March, 2006.
- [2] Huifang Kong and Chris Phillips, “Prebooking for Reduced Reservation Overhead in Long Haul Wavelength-Division Multiplexed (WDM) Networks”, IEE Proceeding of Communications, accepted for publication, Feb, 2006.
- [3] Huifang Kong and Chris Phillips, “Prediction and Pre-booking for Guaranteed Service Provisioning in Future Wavelength-Division Multiplexing (WDM) Networks”, the 19th Internet Teletraffic Congress (ITC’19), Beijing, China, 2005.
- [4] Huifang Kong and Chris Phillips, “A Fast Two-Stage Routing and Wavelength Assignment (RWA) Mechanism for Optimal Max-RWA Problem in Wavelength Constrained WDM Networks”, the 10th European Conference on Networks & Optical Communications (NOC 2005), University College London (UCL), London, UK, July, 2005.
- [5] Huifang Kong and Chris Phillips, “Pre-booking for Latency-Sensitive Service Provisioning in Future Long Haul Wavelength-Division Multiplexing (WDM) Networks with Dynamic Lightpath”, the 10th European Conference on Networks & Optical Communications (NOC 2005), University College London (UCL), London, UK, July, 2005.
- [6] Huifang Kong and Chris Phillips, “A Two-Stage Routing and Wavelength Assignment Mechanism for Paths Calculations in Lambda Switched GMPLS Networks”, Proceeding of IV Workshop in MPLS/GMPLS networks, Girona, Spain, April 2005.
- [7] Huifang Kong and Chris Phillips, “Hybrid Resource Reservation Mechanism for Optical Burst Switched Networks”, Proceeding of SECOND INTERNATIONAL WORKING CONFERENCE, HET-NETs ’04, Ilkley, West Yorkshire, U.K, July 2004.
- [8] Huifang Kong and Chris Phillips, “Intelligent Pre-Booking for Optical Burst Switching”, Proceeding of the International Network Optimization Conference – INOC '2003, Paris (Evry), France, October 2003.
- [9] Huifang Kong and Chris Phillips, “Pre-booking for Optical Burst Switching”, Proceeding of the London Communications Symposium 2003 (LCS), London, UK, September 2003.

Appendix B – Network Topologies

1. Network Topology of USA NSFnet (fiber length in kilometer):



2. Network Topology of European EONnet:



Appendix C – Traffic Matrices for RWA Study

1. Traffic matrix for NSFnet, containing 268 lightpath demands in total:

(Note: the node index is the blue number associated with each node in topology; the vertical node index refers to the source node index, and the horizon node index is the destination node index.)

Node index	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	0	1	3	1	1	1	3	0	2	0	1	2	0	3
1	0	0	0	2	2	2	1	1	1	2	1	0	1	3
2	3	2	0	3	0	1	2	3	1	2	1	2	2	0
3	3	1	0	0	1	1	2	3	2	2	1	2	1	3
4	1	3	0	2	0	1	0	2	0	3	0	1	1	3
5	1	2	1	3	2	0	1	3	3	1	0	1	1	2
6	2	2	3	1	3	3	0	0	3	1	2	0	3	3
7	3	1	2	3	1	0	1	0	0	3	2	0	3	0
8	3	0	1	3	3	3	1	0	0	2	1	1	1	0
9	0	0	0	1	2	0	2	0	1	0	1	0	0	3
10	1	0	0	2	0	3	0	1	0	3	0	3	1	3
11	2	3	1	1	3	2	3	2	2	2	2	0	1	3
12	2	0	1	2	0	1	2	0	3	0	2	1	0	3
13	1	1	0	2	1	0	1	3	0	1	2	1	3	0

2. Traffic matrix for EONnet, containing 374 lightpath demands in total:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	0	1	2	1	1	0	2	0	1	0	1	2	0	2	1	0	1	1	1	0
1	1	0	0	2	0	0	1	2	2	1	2	0	1	1	2	0	2	0	1	1
2	0	2	0	0	0	1	1	1	2	1	1	1	1	0	2	1	2	0	1	0
3	0	1	0	0	2	0	0	0	2	1	2	0	2	2	1	2	2	1	0	1
4	0	2	2	1	0	2	1	2	2	0	2	1	1	0	2	2	2	1	2	2
5	1	0	1	0	2	0	1	0	2	0	2	0	0	2	2	2	1	0	1	0
6	0	0	0	0	0	0	0	1	2	0	1	0	1	1	0	0	2	1	0	0
7	1	0	2	0	1	0	2	0	2	1	2	2	2	1	1	2	2	2	2	1
8	2	1	0	2	1	0	1	1	0	0	1	1	0	2	0	2	0	2	1	0
9	0	1	0	0	0	2	0	0	1	0	0	2	0	2	2	2	1	0	2	0
10	1	2	2	1	2	0	2	1	2	1	0	2	1	2	2	0	2	0	1	0
11	1	1	0	1	1	2	1	0	1	0	0	0	0	0	2	1	0	2	0	0
12	2	2	2	2	0	0	1	1	1	0	1	2	0	0	0	1	1	0	2	1
13	0	0	2	2	0	2	0	1	2	1	2	1	1	0	2	1	1	0	0	1
14	1	0	2	0	1	0	0	1	0	2	2	2	0	2	0	2	2	1	2	1
15	1	0	1	0	1	1	2	0	0	2	2	0	1	1	2	0	1	2	1	2
16	0	0	1	2	2	1	1	2	0	0	1	2	0	2	2	1	0	1	1	1
17	0	1	2	0	2	2	2	0	1	2	2	0	2	1	0	1	0	0	2	0
18	1	0	1	0	2	2	1	0	2	1	2	1	0	2	0	1	1	1	0	2
19	1	2	2	0	1	0	0	0	0	1	0	0	0	2	2	0	1	2	2	0

Appendix D - Mathematical Models for Standard JET System and Pre-booking Mechanism

1. Work-conserving Model for Standard JET System

At each OBS node with k outgoing wavelength channels, if incoming bursts follow a Poisson arrival process with mean arrival rate λ and burst length is exponentially distributed, the traffic intensity is given by Equation D-1, where μ is the mean service rate (pkts/sec), and $r=\rho \times k$.

$$\rho = \lambda / (\mu \cdot k) = r / k \quad \text{Equation D-1}$$

Such an OBS node behaves like an M/M/k/k loss system for which burst blocking probability can be obtained using the *Erlang B formula* as in Equation D-2 [YQ99].

$$P = B(k, \rho) = (1 / k! \cdot r^k) / \left(\sum_{m=0}^k 1 / m! \cdot r^m \right) \quad \text{Equation D-2}$$

2. Work-conserving Model for Simple Pre-booking System

In the proposed pre-booking system, there are two groups of independent traffic, predicted and non-predicted traffic. The predicted traffic intensity is denoted by ρ_0 , and the non-predicted traffic intensity is denoted by ρ_1 . The total traffic intensity is $\rho_{all}=\rho_0+\rho_1$.

As the predicted traffic is completely independent of the non-predicted traffic, and is able to pre-empt resources before the non-predicted traffic comes up, the system with the predicted traffic alone is an M/M/k/k system. The blocking probability for predicted traffic is $P_0=B(k, \rho_0)$.

Assuming the whole system still behaves as an M/M/k/k loss system regardless of the degree of traffic isolation, the overall blocking probability can also be obtained by $P_{all}=B(k, \rho_{all})$.

Thus, the blocking probability for non-predicted traffic can be obtained using the *Conservation Law* as shown in Equation D-3 [YQ99][HM02]:

$$\rho_{all} P_{all} = \rho_0 P_0 + \rho_1 P_1 \quad \text{Equation D-3}$$

Where P_{all} is the overall blocking probability, P_0 is the predicted traffic blocking probability, and P_1 is the non-predicted traffic blocking probability.

Appendix E – Validation of Simulation Models for Pre-booking Mechanism in Classical OBS Networks

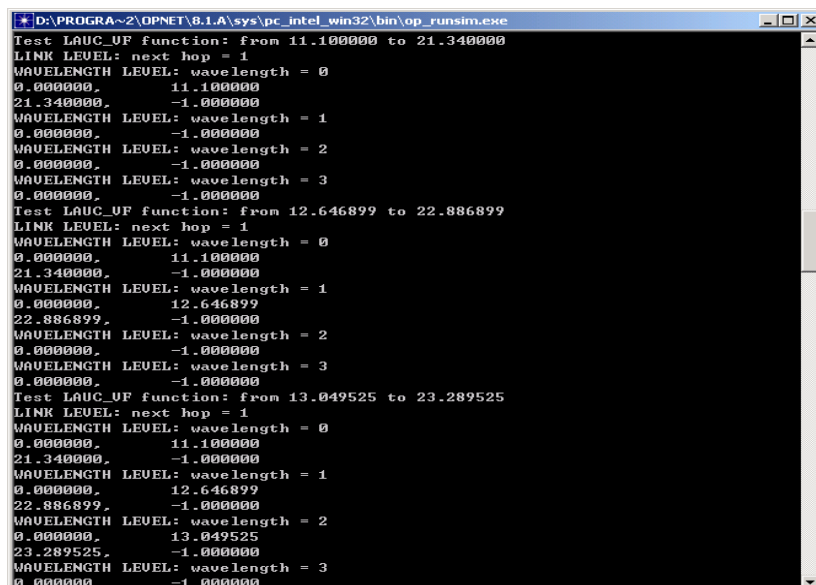
This appendix provides description and snapshots for validating the pre-booking simulation models in classical OBS networks. The validation for pre-booking models in hybrid network architecture and WR-OBS based architectures adopt the same methodology and procedure as the material provided in this appendix.

1. Preliminary Model with Packet Transmission and Receiving

The basic validation methodology for this stage is to trace the bursts and the corresponding BHP by printing out their status. The printed messages confirm that the simulator deals the events properly and does not mis-sequence or arbitrary abandon any events.

1.1 The LAUC-VF Algorithm

The LAUC-VF algorithm [XVC00] is adopted in JET to manage the resource reservation procedure. For validation, two tests were carried out. The first test was based on a constant packet arrival rate. The second test validates the algorithm with a stochastic packet arrival rate to confirm the satisfactory operation of a more general form of the model. The parameters for the second test were: Wavelength number = 4; Packet Inter-arrival Time = Exponential Distribution with Mean of 1.0 sec.; Burst Size = 102400 bits; and Wavelength Rate = 10 kbits/sec. The snapshots for the second test are shown as following, in which all the available wavelength time slots on the link concerned are monitored for each reservation request.



```
D:\PROGRA~2\OPNET\6.1.A\sys\pc_intel_win32\bin\op_runsim.exe
Test LAUC_VF function: from 11.100000 to 21.340000
LINK LEVEL: next hop = 1
WAULENGTH LEVEL: wavelength = 0
0.000000, 11.100000
21.340000, -1.000000
WAULENGTH LEVEL: wavelength = 1
0.000000, -1.000000
WAULENGTH LEVEL: wavelength = 2
0.000000, -1.000000
WAULENGTH LEVEL: wavelength = 3
0.000000, -1.000000
Test LAUC_VF function: from 12.646899 to 22.886899
LINK LEVEL: next hop = 1
WAULENGTH LEVEL: wavelength = 0
0.000000, 11.100000
21.340000, -1.000000
WAULENGTH LEVEL: wavelength = 1
0.000000, 12.646899
22.886899, -1.000000
WAULENGTH LEVEL: wavelength = 2
0.000000, -1.000000
WAULENGTH LEVEL: wavelength = 3
0.000000, -1.000000
Test LAUC_VF function: from 13.049525 to 23.289525
LINK LEVEL: next hop = 1
WAULENGTH LEVEL: wavelength = 0
0.000000, 11.100000
21.340000, -1.000000
WAULENGTH LEVEL: wavelength = 1
0.000000, 12.646899
22.886899, -1.000000
WAULENGTH LEVEL: wavelength = 2
0.000000, 13.049525
23.289525, -1.000000
WAULENGTH LEVEL: wavelength = 3
0.000000, -1.000000
```

```

D:\PROGRA~2\OPNET\6.1.A\sys\pc_intel_win32\bin\op_runsim.exe
Test LAUC_VF function: from 13.170209 to 23.410209
LINK LEVEL: next hop = 1
WAVELENGTH LEVEL: wavelength = 0
0.000000, 11.100000
21.340000, -1.000000
WAVELENGTH LEVEL: wavelength = 1
0.000000, 12.646899
22.886899, -1.000000
WAVELENGTH LEVEL: wavelength = 2
0.000000, 13.049525
23.289525, -1.000000
WAVELENGTH LEVEL: wavelength = 3
0.000000, 13.170209
23.410209, -1.000000
Test LAUC_VF function: from 13.221782 to 23.461782
LINK LEVEL: next hop = 1
WAVELENGTH LEVEL: wavelength = 0
0.000000, 11.100000
21.340000, -1.000000
WAVELENGTH LEVEL: wavelength = 1
0.000000, 12.646899
22.886899, -1.000000
WAVELENGTH LEVEL: wavelength = 2
0.000000, 13.049525
23.289525, -1.000000
WAVELENGTH LEVEL: wavelength = 3
0.000000, 13.170209
23.410209, -1.000000

```

```

D:\PROGRA~2\OPNET\6.1.A\sys\pc_intel_win32\bin\op_runsim.exe
Test LAUC_VF function: from 13.491838 to 23.731838
LINK LEVEL: next hop = 1
WAVELENGTH LEVEL: wavelength = 0
0.000000, 11.100000
21.340000, -1.000000
WAVELENGTH LEVEL: wavelength = 1
0.000000, 12.646899
22.886899, -1.000000
WAVELENGTH LEVEL: wavelength = 2
0.000000, 13.049525
23.289525, -1.000000
WAVELENGTH LEVEL: wavelength = 3
0.000000, 13.170209
23.410209, -1.000000
Test LAUC_VF function: from 14.805116 to 25.045116
LINK LEVEL: next hop = 1
WAVELENGTH LEVEL: wavelength = 0
0.000000, 11.100000
21.340000, -1.000000
WAVELENGTH LEVEL: wavelength = 1
0.000000, 12.646899
22.886899, -1.000000
WAVELENGTH LEVEL: wavelength = 2
0.000000, 13.049525
23.289525, -1.000000
WAVELENGTH LEVEL: wavelength = 3
0.000000, 13.170209
23.410209, -1.000000
:
: Simulation Completed - Collating Results.
: Events: Total (29), Average Speed (1449 events/sec.)
: Time: Elapsed (1 sec.), Simulated (14 sec.)
:
Press <ENTER> to continue.

```

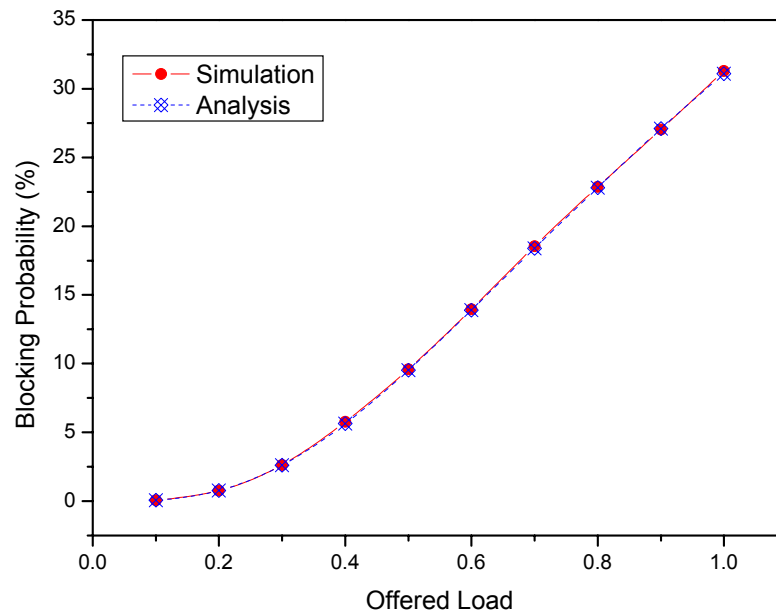
Snapshots for testing the LAUC-VF algorithm

1.2 The JET Model

As the standard JET model can be theoretically modelled as an M/M/k/k loss system with Poisson incoming traffic, the validation is done by comparing the simulated blocking probability with analytical blocking rate.

The figure below shows the blocking probability for both simulation and analytical models. Every point of each simulation result is actually the average of ten independent experiments with ten different seeds, such as to get more stable and reliable results with a reasonable confidence level. The analytical results are drawn from the Erlang B formula for an M/M/k/k loss system. As the simulation results are consistent with the analysis results, the JET model is

considered to be valid.



Validation of JET Model

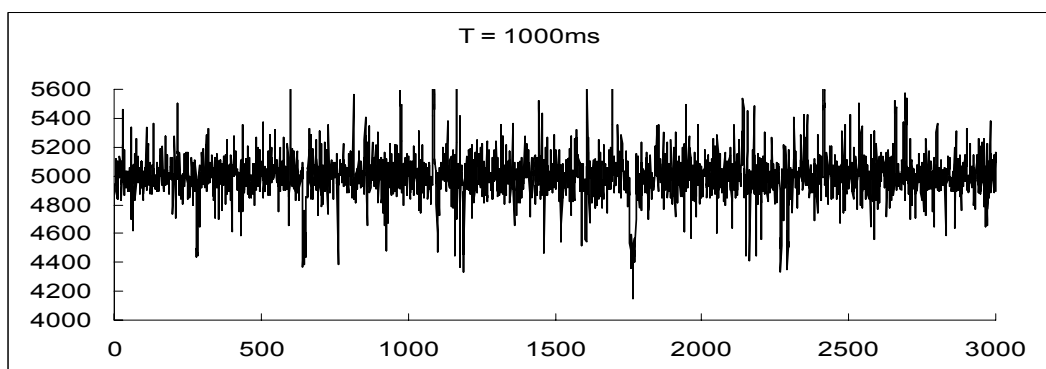
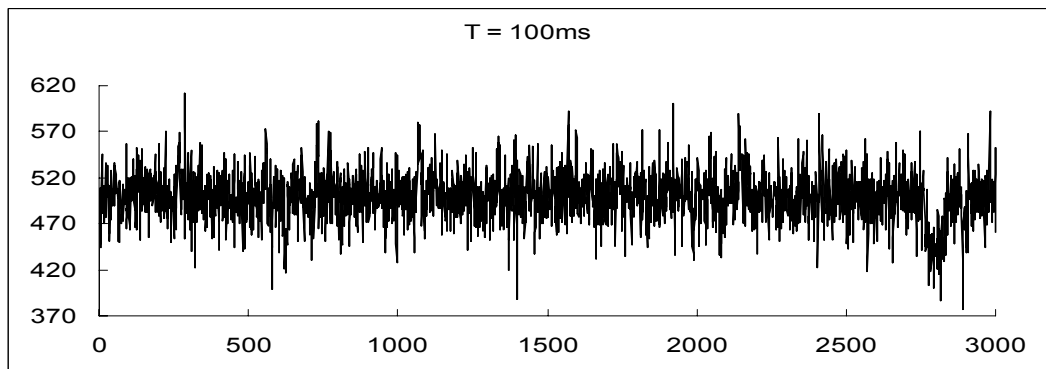
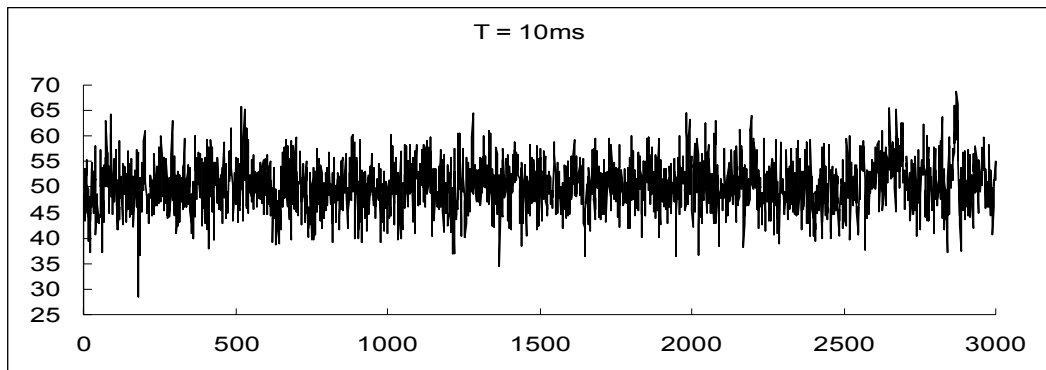
1.3 The Simple Pre-booking Model

Compared with JET model, the simple pre-booking case only differs in the source model where there are two groups of traffic, predicted and non-predicted traffic. In terms of validation, the same methodology as in JET model is adopted. This is to cross-compare the simulation and analytical results. As presented in Fig 4-10, the blocking probability of analysis and simulation results match each other very well. Therefore, the model appears correct.

Appendix F - Validating the Self-Similar Traffic

This appendix exams the self-similar traffic produced by superposition multiple ON-OFF processes, and points out advantages and limitations associated with the model.

Figure below shows samples of traffic generated from superposition model with 10 ON-OFF sources.

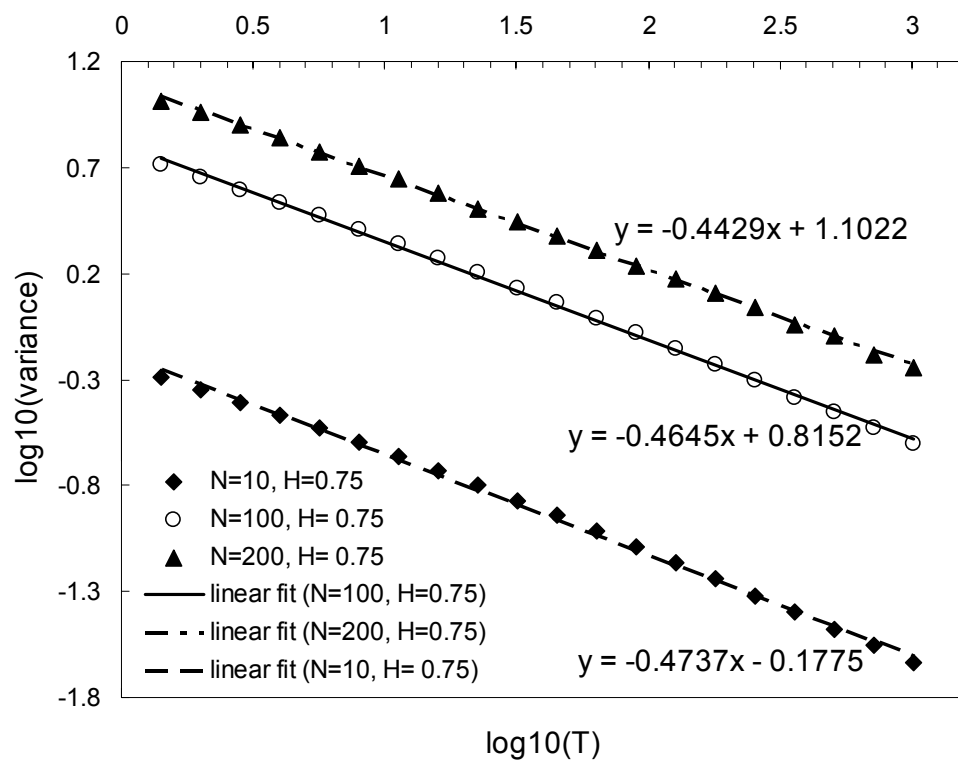


Traffic samples from 10 ON-OFF processes

Each ON-OFF model yields the same Pareto distribution for ON and OFF periods. The location

parameter $k = 0.0352$, which is equivalent to the time (in milliseconds) required to transmit a 44Byte packet with 10Mbps transmission rate on each stream [XVC00]. The shape parameter $\alpha = 1.5$, leading to a theoretical *Hurst* parameter $H = 0.75$. The figure above demonstrates that the generated traffic yields similar shapes and spiky characteristics at different levels of aggregation, which is a sign of self-similarity.

The self-similarity of the generated traffic should be strictly validated via the variance-time plot, which is shown in figure blow. The plot shows results from three models with 10, 100, and 200 ON-OFF processes, respectively.



Variance-time plot for traffic from superposition of ON-OFF processes

Figure above demonstrates that there is an approximately linear relationship between the logarithmic aggregation time and the logarithmic aggregated burst length for all the three models. However, the slopes of the three models differ from each other. For example, the theoretical slope is -0.5 ($slope = 2(H-1)$), whilst, the slope with 10 ON-OFF processes is about -0.4737 , the slopes for the other two models are -0.4645 , and -0.4429 , respectively. Though the slopes are close to each other, they deviate from the theoretical slope with unpredictable gaps,

where the gap may be caused by the finite number of sources.

In a simulation environment, it is very hard to determine the optimum number of ON-OFF processes such that the resulted *Hurst* parameter is as close to the theoretical value as possible. Additionally, it is also very hard to predict the variance of the generated traffic. The only way to obtain the variance is to do the calculation based on the whole data set collected from the simulation.

References

- [Adas98] Adas, A.M.; “Using adaptive linear prediction to support real-time VBR video under RCBR network service model”, IEEE/ACM Transactions on Networking, Volume: 6, Issue: 5, Oct. 1998 Pages:635 – 644.
- [Alc1626] Alcatel 1626 LM Platform: <http://www.alcatel.com/products/productssummary.jhtml?repositoryID=/com/en/appxml/opgproduct/alcatel1626lightmanagertcm228119011635.jhtml>.
- [AMD02] Erramilli, A.; Roughan, M.; Veitch, D.; Willinger, W.; “Self-similar traffic and network dynamics”, IEEE Proceedings, Volume 90, Issue 5, May 2002, Page(s):800 – 819.
- [APP01] Andre, P.S.; Pinto, A.N.; Pinto, J.L.; Almeida, T.; Pousa, M., “Tunable transparent and cost effective optical add-drop multiplexer based on fiber Bragg grating for DWDM networks”, Advanced Semiconductor Lasers and Applications/Ultraviolet and Blue Lasers and Their Applications/Ultralong Haul DWDM Transmission and Networking/WDM Components, 2001 Digest of the LEOS Summer Topical Meetings, 30 July-1 Aug. 2001 Page(s):2 pp.
- [BB97] Baroni, S.; Bayvel, P., “Wavelength requirements in arbitrarily connected wavelength-routed optical networks”, Journal of Lightwave Technology, Volume 15, Issue 2, Feb. 1997 Page(s):242 – 251.
- [BCB04] Jaumard, B.; Meyer, C.; Thiongane, B.; “ILP formulations and optimal solutions for the RWA problem”, IEEE Global Telecommunications Conference, 2004. GLOBECOM '04, Volume 3, 29 Nov.-3 Dec. 2004 Page(s):1918 - 1924 Vol.3.
- [BJH02] Di Blas, A.; Jagota, A.; Hughey, R.; “Energy function-based approaches to graph coloring”, IEEE Transactions on Neural Networks, Volume 13, Issue 1, Jan. 2002 Page(s):81 – 91, Digital Object Identifier 10.1109/72.977273.
- [BM96] Banerjee, D.; Mukherjee, B., “A Practical Approach for Routing and Wavelength Assignment in Large Wavelength-Routed Optical Networks”, IEEE Journal on Selected Areas in Communications, Volume: 14, Issue: 5, June 1996, Pages: 903 – 908.
- [BOR00] Blumenthal, D.J.; Olsson, B.-E.; Rossi, G.; Dimmick, T.E.; Rau, L.; Masanovic, M.; Lavrova, O.; Doshi, R.; Jerphagnon, O.; Bowers, J.E.; Kaman, V.; Coldren, L.A.; Barton, J., “All-optical label swapping networks and technologies”, Journal of Lightwave Technology, Volume 18, Issue 12, Dec 2000, Page(s):2058 – 2075.
- [CCL97] Huimin Chen; Hong Cai; Yanda Li; “The multifractal property of bursty traffic and its parameter estimation based on wavelets”, IEEE Region 10 Annual Conference TENCON '97. Proceedings of IEEE Speech and Image Technologies for Computing and Telecommunications', Volume 2, 2-4 Dec. 1997 Page(s):791 - 794 vol.2.
- [CFC03] Chuck Fraleigh, Fouad Tobagi, and Christophe Diot.; “Provisioning IP backbone networks to support latency sensitive traffic”. In Proceedings of 22nd Annual Joint

Conference of the IEEE Computer and Communications Societies (INFOCOM), March/April 2003.

- [CGG01] M. Corradi, R. G. Garroppo, S. Giordana and M. Pagano, "Analysis of F-ARIMA processes in the modelling of broadband traffic", ICC'01, vol. 3, pp. 964-968, 2001.
- [Che92] Cheung, N.K., "The infrastructure for gigabit computer networks", IEEE Communications Magazine, Volume 30, Issue 4, April, 1992, Page(s):60 – 68.
- [CKS04] Cavendish, D.; Kolarov, A.; Sengupta, B., "Routing and wavelength assignment in WDM mesh networks", IEEE Global Telecommunications Conference, 2004. GLOBECOM '04, Volume 2, 29 Nov.-3 Dec. 2004, Page(s):1016 - 1022 Vol.2.
- [CL96] Joseph C. Culberson and Feng Luo, "Exploring the k-colorable Landscape with Iterated Greedy", DIMACS Series for "Cliques, Coloring and Satisfiability", Volume 26, 245-284, 1996.
- [CL98] Yung-Kuang Chen; Chien-Chung Lee, "Fiber Bragg grating-based large nonblocking multiwavelength cross-connects", Journal of Lightwave Technology, Volume 16, Issue 10, Oct. 1998, Page(s):1746 – 1756.
- [CO02] K.G. Coffman and A.M. Odlyzko, "Growth of the Internet", in Optical Fiber Communications, vol. IV-B: Systems and Impairments, I.P. Kaminow and T. Li(ed), Academic Press, San Diego, 2002, pp. 17-56.
- [Cav00] Cavendish, D.; "Evolution of optical transport technologies: from SONET/SDH to WDM", IEEE Communications Magazine, Volume 38, Issue 6, June 2000, Page(s):164 – 172.
- [CJ05] Pinart, C.; Giralt, G.J.; "On managing optical services in future control-plane-enabled IP/WDM networks", Journal of Lightwave Technology, Volume 23, Issue 10, Oct. 2005 Page(s):2868 – 2876.
- [CQY04] Yang Chen; Chunming Qiao; Xiang Yu; "Optical burst switching: a new area in optical networking research", IEEE Network, Volume 18, Issue 3, May-June 2004, Page(s):16 – 23.
- [Daniel79] Daniel Brelaz, "New methods to color the vertices of a graph", Communications of the ACM, Volume 22, Issue 4, April 1979, Pages: 251 – 256, ISSN:0001-0782.
- [DB02] Duser, M.; Bayvel, P., "Analysis of a dynamically wavelength-routed optical burst switched network architecture", Journal of Lightwave Technology, Volume 20, Issue 4, April 2002, Page(s):574 – 585.
- [DB022] Duser, M.; Bayvel, P.; "Performance of a dynamically wavelength-routed optical burst switched network", IEEE Photonics Technology Letters, Volume 14, Issue 2, Feb. 2002, Page(s):239 – 241.
- [DGS01] K. Dolzer, C. Gauger, J. Spath, and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching", AEU International Journal of Electronics and Communications, 55 (1), January 2001.
- [DHL03] Dorren, H.J.S.; Hill, M.T.; Liu, Y.; Calabretta, N.; Srivatsa, A.; Huijskens, F.M.; de

- Wardt, H.; Khoe, G.D., "Optical packet switching and buffering by using all-optical signal processing methods", *Journal of Lightwave Technology*, Volume 21, Issue 1, Jan 2003, Page(s):2 – 12.
- [DJP02] Doyle, A.; Juignet, C.; Painchaud, Y.; Brown, A.; Chummun-Courbet, N.; Pelletier, T.; Guy, A., "FBG-based multi-channel low dispersion WDM filters", *Electronics Letters*, Volume 38, Issue 24, 21 Nov. 2002, Page(s):1561 – 1563.
- [Duser03] Michael Duser, "Investigation of Advanced Optical Packet-Routed Network Architectures", PhD Thesis, Department of Electronic and Electrical Engineering, University College London, Jan 2003.
- [DW99] Daniel, M.M.; Willsky, A.S.; "The modeling and estimation of statistically self-similar processes in a multiresolution framework", *IEEE Transactions on Information Theory*, Volume 45, Issue 3, April 1999, Page(s):955 – 970.
- [DZB04] Michael Düser, Alejandra Zapata, and Polina Bayvel; "Investigation of the scalability of dynamic wavelength-routed optical networks [Invited]", *Journal of Optical Networking*, September 2004, Volume: 3, No. 9, August 2004, Pages:667 – 686.
- [EA97] Crovella, M.E.; Bestavros, A., "Self-similarity in World Wide Web traffic: evidence and possible causes", *IEEE/ACM Transactions on Networking*, Volume 5, Issue 6, Dec. 1997 Page(s):835 – 846.
- [ESG01] Varma, E.L.; Sankaranarayanan, S.; Newsome, G.; Zhi-Wei Lin; Epstein, H.; "Architecting the services optical network", *IEEE Communications Magazine*, Volume 39, Issue 9, Sept. 2001, Page(s):80 – 87.
- [ESW94] Leland, W.E.; Taqqu, M.S.; Willinger, W.; Wilson, D.V.; "On the self-similar nature of Ethernet traffic (extended version)", *IEEE/ACM Transactions on Networking*, Volume 2, Issue 1, Feb. 1994, Page(s):1 – 15.
- [FBS00] Fabry-Asztalos, T.; Bhide, N.; Sivalingam, K.M., "Adaptive weight functions for shortest path routing algorithms for multi-wavelength optical WDM networks", *IEEE International Conference on Communications, 2000 (ICC 2000)*, Volume 3, 18-22, June 2000, Page(s):1330 - 1334 vol.3.
- [GCT00] Ge, A.; Callegati, F.; Tamil, L.S., "On optical burst switching and self-similar traffic", *IEEE Communications Letters*, Volume 4, Issue 3, March 2000, Page(s):98 – 100.
- [GJ79] M. R. Garey and D. S. Johnson, "Computers and Intractability: A Guide to the Theory of NP-Completeness". New York, NY: Freeman, 1979.
- [GP94] Groschwitz, N.K.; Polyzos, G.C.; "A time series model of long-term NSFNET backbone traffic", *IEEE International Conference on Communications, 1994. ICC 94, SUPERCOMM/ICC '94, Conference Record, Serving Humanity Through Communications, 1-5 May 1994* Pages:1400 - 1404 vol.3.
- [GRS00] Gerstel, O.; Ramaswami, R.; Sasaki, G.H., "Cost-effective traffic grooming in WDM rings", *IEEE/ACM Transactions on Networking*, Volume 8, Issue 5, Oct. 2000 Page(s):618 – 630.

- [GRG98] Gambini, P.; Renaud, M.; Guillemot, C.; Callegati, F.; Andonovic, I.; Bostica, B.; Chiaroni, D.; Corazza, G.; Danielsen, S.L.; Gravey, P.; Hansen, P.B.; Henry, M.; Janz, C.; Kloch, A.; Krahenbuhl, R.; Raffaelli, C.; Schilling, M.; Talneau, A.; Zucchelli, L., "Transparent optical packet switching: network architecture and demonstrators in the KEOPS project", IEEE Journal on Selected Areas in Communications, Volume 16, Issue 7, Sept. 1998, Page(s):1245 – 1259.
- [GSP94] Glesk, I.; Solokoff, J.P.; Prucnal, P.R., "All-optical address recognition and self-routing in a 250 Gbit/s packet-switched network", Electronics Letters, Volume 30, Issue 16, 4 Aug. 1994 Page(s):1322 – 1323.
- [GW99] Gilbert, A.C.; Willinger, W.; Feldmann, A.; "Scaling analysis of conservative cascades, with applications to network traffic", IEEE Transactions on Information Theory, Volume 45, Issue 3, April 1999 Page(s):971 – 991.
- [HDG03] Hu, G.; Dolzer, K.; Gauger, C., "Does burst assembly really reduce the self-similarity?", Optical Fiber Communications Conference, 2003. OFC 2003, 23-28 March 2003 Page(s):124 - 126 vol.1.
- [HM02] H. Vu and M. Zukerman. "Blocking probability for priority classes in optical burst switching networks". IEEE Communications Letters, 6:214-216, May 2002.
- [HYH98] Kun-Yuan Huang; Hong-Tzer Yang; Ching-Lien Huang, "A new thermal unit commitment approach using constraint logic programming", IEEE Transactions on Power Systems, Volume 13, Issue 3, Aug. 1998, Page(s):936 – 945.
- [IA02] Izal, M.; Aracil, J., "On the influence of self-similarity on optical burst switching traffic", IEEE Global Telecommunications Conference, 2002. GLOBECOM '02, Volume 3, 17-21 Nov. 2002 Page(s):2308 - 2312 vol.3.
- [JMT04] Jaumard, B.; Meyer, C.; Thiongane, B., "ILP formulations and optimal solutions for the RWA problem", IEEE Global Telecommunications Conference, 2004. GLOBECOM '04, Volume 3, 29 Nov.-3 Dec. 2004 Page(s):1918 - 1924 Vol.3.
- [JNM03] Kuri, J.; Puech, N.; Gagnaire, M.; Dotaro, E.; Douville, R., "Routing and wavelength assignment of scheduled lightpath demands", IEEE Journal on Selected Areas in Communications, Volume: 21, Issue: 8, Oct. 2003, Pages:1231 - 1240.
- [KB02] E. Kozlovski and P. Bayvel, "QoS performance of WR-OBS network architecture with request scheduling", Proceeding of IFIP 6th Working Conference on Optical Networks Design and Modelling (ONDM 2002). Turin, Italy. February 2002.
- [KDZ02] E. Kozlovski, M. Dueser, A. Zapata and P. Bayvel, "Service differentiation in wavelength-routed optical burst-switched networks", Proceeding of IEEE/OSA Conference on Optical Fibre Communications (OFC2002). Anaheim, USA. pp.774-775. March 2002.
- [KKK02] Sungchang Kim; Namook Kim; Minho Kang; "Contention resolution for optical burst switching networks using alternative routing", IEEE International Conference on Communications, 2002. ICC 2002, Volume: 5, 28 April-2 May 2002 Page(s): 2678 -2681 vol.5.

- [KS01] Krishnaswamy, R.M.; Sivarajan, K.N., “Algorithms for Routing and Wavelength Assignment Based on Solutions of LP-Relaxations”, IEEE Communications Letters, Volume: 5, Issue: 10, Oct. 2001, Pages : 435 – 437.
- [LAO03] Jingxuan Liu, N. Ansari, and T.J. Ott, “FRR for latency reduction and QoS provisioning in OBS networks”, in Selected Areas in Communications, IEEE Journal on Volume 21, Issue 7, Sept. 2003 Page(s):1210 – 1219.
- [Liang04] Yao Liang; “Real-time VBR video traffic prediction for dynamic bandwidth allocation”, IEEE Transactions on Systems, Man and Cybernetics, Part C, Volume 34, Issue 1, Feb. 2004, Page(s):32 – 47.
- [LN02] Jingxuan Liu; Ansari, N.; “Forward resource reservation for QoS provisioning in OBS systems”, IEEE Global Telecommunications Conference, 2002. GLOBECOM '02, Volume 3, 17-21 Nov. 2002 Page(s):2777 - 2781 vol.3.
- [LN03] Jingxuan Liu; Ansari, N.; “A bandwidth enhancement mechanism for the FRR-enabled OBS networks”, Workshop on High Performance Switching and Routing, 2003, HPSR., 24-27 June 2003 Page(s):135 – 139.
- [LNJ03] Jingxuan Liu; Ansari, N.; Ott, T.J.; “FRR for latency reduction and QoS provisioning in OBS networks”, IEEE Journal on Selected Areas in Communications, Volume 21, Issue 7, Sept. 2003 Page(s):1210 – 1219.
- [LT03] Lin, P.J.; Narula-Tam, A., “Cell-sorting device for creating synchronous variable-length optical packet switches”, Journal of Lightwave Technology, Volume 21, Issue 4, April 2003, Page(s):893 – 903.
- [MDB01] I. de Miguel, M. Dueser, and P. Bayvel, “Traffic Load Bounds for Optical Burst-Switched Networks with Dynamic Wavelength Allocation”, Proceeding of IFIP 5th Working Conference on Optical Network Design and Modelling (ONDM 2001), vol. 1, Vienna, 5 - 7 Feb. 2001.
- [MK03] Nam Su Moon; Kikuchi, K., “N×N multiwavelength optical cross-connect based on tunable fiber Bragg gratings”, Journal of Lightwave Technology, Volume 21, Issue 3, March 2003, Page(s):703 – 718.
- [MKB02] I. de Miguel, E. Kozlovski and P. Bayvel, “Provision of end-to-end delay guarantees in wavelength-routed optical burst-switched networks”, in Proc. IFIP 6th Working Conference on Optical Networks Design and Modelling (ONDM 2002). Turin, Italy. February 2002.
- [MNG05] Marom, D.M.; Neilson, D.T.; Greywall, D.S.; Chien-Shing Pai; Basavanhally, N.R.; Aksyuk, V.A.; Lopez, D.O.; Pardo, F.; Simon, M.E.; Low, Y.; Kolodner, P.; Bolle, C.A., “Wavelength-selective 1/spl times/K switches using free-space optics and MEMS micromirrors: theory, design, and implementation”, Journal of Lightwave Technology, Volume 23, Issue 4, April 2005, Page(s):1620 – 1630.
- [MT98] M. Matsumoto and T. Nishimura, “Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator”, ACM Trans. on Modeling and Computer Simulation Vol. 8, No. 1, January pp.3-30 (1998).

- [NRH03] Neuts, M.; Rosberg, Z.; Hai Le Vu; White, J.; Zukerman, M.; “Performance enhancement of optical burst switching using burst segmentation”, IEEE International Conference on Communications, 2003. ICC '03, Volume: 3 , 11-15 May 2003 Page(s): 1828 -1832.
- [Pax95] V. Paxson, “Fast Approximation of Self-Similar Network Traffic”, Lawrence Berkeley National Lab., Tech. Rep. LBL-36 750, Apr. 1995.
- [PGP93] Narain, P.; Saab, D.G.; Kunda, R.P.; Abraham, J.A.; “A high-level approach to test generation”, IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, [see also IEEE Transactions on Circuits and Systems I: Regular Papers,], Volume 40, Issue 7, July 1993 Page(s):483 – 492.
- [Qiao00] Chunming Qiao; “Labeled optical burst switching for IP-over-WDM integration”, IEEE Communications Magazine, Volume 38, Issue 9, Sept. 2000 Page(s):104 – 114.
- [RFC3031] RFC 3031 - Multiprotocol Label Switching Architecture.
- [RFC3473] RFC 3473 - Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions.
- [RKH01] Ryf, R.; Kim, J.; Hickey, J.P.; Gnauck, A.; Carr, D.; Pardo, F.; Bolle, C.; Frahm, R.; Basavanhally, N.; Yoh, C.; Ramsey, D.; Boie, R.; George, R.; Kraus, J.; Lichtenwalner, C.; Papazian, R.; Gates, J.; Shea, H.R.; Gasparyan, A.; Muratov, V.; Griffith, J.E.; Prybyla, J.A.; Goyal, S.; White, C.D.; Lin, M.T.; Ruel, R.; Nijander, C.; Arney, S.; Neilson, D.T.; Bishop, D.J.; Kolodner, P.; Pau, S.; Nuzman, C.; Weis, A.; Kumar, B.; Lieuwen, D.; Aksyuk, V.; Greywall, D.S.; Lee, T.C.; Soh, H.T.; Mansfield, W.M.; Jin, S.; Lai, W.Y.; Huggins, H.A.; Barr, D.L.; Cirelli, R.A.; Bogart, G.R.; Teffeau, K.; Vella, R.; Mavoori, H.; Ramirez, A.; Ciampa, N.A.; Klemens, F.P.; Morris, M.D.; Boone, T.; Liu, J.Q.; Rosamilia, J.M.; Giles, C.R., “1296-port MEMS transparent optical crossconnect with 2.07 petabit/s switch capacity”, Optical Fiber Communication Conference and Exhibit, 2001 (OFC 2001). Volume 4, 2001 Page(s):PD28-1 - PD28-3 vol.4.
- [Rob01] Robinson, S.D., “MEMS technology-micromachines enabling the all optical network”, Proceedings of 51st Electronic Components and Technology Conference, 2001, 29 May-1 June 2001, Page(s):423 – 428.
- [ROB04] Rajaduray, R.; Ovadia, S.; Blumenthal, D.J.; “Analysis of an edge router for span-constrained optical burst switched (OBS) networks”, Journal of Lightwave Technology, Volume 22, Issue 11, Nov. 2004 Page(s):2693 – 2705.
- [RS95] Ramaswami, R.; Sivarajan, K.N., “Routing and Wavelength Assignment in All-Optical Networks”, IEEE/ACM Transactions on Networking, Volume: 3, Issue: 5, Oct. 1995, Pages: 489 – 500.
- [SJZ99] Y. Shu, Z. Jin, L. Zhang, L. Wang and O. W. W. Yang, “Traffic prediction using FARIMA models”, ICC'99, vol. 2, pp. 891-895, 1999.
- [SKC03] N. Sadek, A. Khotanzad, T. Chen, “ATM dynamic bandwidth allocation using F-ARIMA prediction model”, 12th International Conf. on Computer

Communications and Networks (ICCCN 2003), Oct. 20-22, 2003.

- [SLK05] A. Stavdas, H. Leligou, K. Kanonakis, C. Linardakis, and J. D. Angelopoulos, "Scheme for performing statistical multiplexing in the optical layer," *Journal of Optical Networking*. 4, 237-247 (2005).
- [SM04] Sharafat, A.R.; Ma'rouzi, O.R., "The most congested cutset: deriving a tight lower bound for the chromatic number in the RWA problem", *IEEE Communications Letters*, Volume 8, Issue 7, July 2004, Page(s):473 – 475.
- [SST03] Sahara, A.; Shimano, K.; Takigawa, Y.; Koga, M.; "Optical burst data switching utilising GMPLS signaling", *Electronics Letters*, Volume 39, Issue 17, 21 Aug. 2003, Page(s):1267 – 1269.
- [STS04] Sahara, A.; Tsukishima, Y.; Shimano, K.; Koga, M.; Mori, K.; Sakai, Y.; Ishii, Y.; Kawai, M.; "Demonstration of connection-oriented optical burst switching network utilising PLC and MEMS switches", *Electronics Letters*, Volume 40, Issue 25, 9 Dec. 2004 Page(s):1597 – 1599.
- [Sue05] Sue, C.-C.; "Wavelength routing with spare reconfiguration for all-optical WDM networks", *Journal of Lightwave Technology*, Volume 23, Issue 6, June 2005, Page(s):1991 – 2000.
- [Taga96] Taga, H.; "Long distance transmission experiments using the WDM technology", *Journal of Lightwave Technology*, Volume 14, Issue 6, June 1996 Page(s):1287 – 1298.
- [TAK03] Tachibana, T.; Ajima, T.; Kasahara, S., "Round-robin burst assembly and constant transmission scheduling for optical burst switching networks", *IEEE Global Telecommunications Conference, 2003. GLOBECOM '03. Volume 5, 2003* Page(s):2772 - 2776 vol.5.
- [TMP02] Tomatore, M.; Maier, G.; Pattavina, A., "WDM network optimization by ILP based on source formulation", *Proceedings of IEEE INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies*, Volume 3, 23-27 June 2002 Page(s):1813 - 1821 vol.3.
- [TNT93] Iokibe, T.; Mochizuki, N.; Kimura, T.; "Traffic prediction method by fuzzy logic", *Second IEEE International Conference on Fuzzy Systems*, 1993, 28 March-1, April 1993 Page(s):673 - 678 vol.2.
- [TSP00] Tang, J.M.; Spencer, P.S.; Rees, P.; Shore, K.A., "Ultrafast optical packet switching using low optical pulse energies in a self-synchronization scheme", *Journal of Lightwave Technology*, Volume 18, Issue 12, Dec 2000, Page(s):1757 – 1764.
- [TZT03] Tzanakaki, A.; Zacharopoulos, I.; Tomkos, I., "Optical add/drop multiplexers and optical cross-connects for wavelength routed networks", *Proceedings of 2003 5th International Conference on Transparent Optical Networks*, 2003. Volume 1, 29 June-3 July 2003, Page(s):41 - 46 vol.1.
- [UE04] Uengsakul, J.; Erke, T., "Performance evaluation of assembly mechanisms for jet-based optical burst switching networks", *IEEE Region 10 Conference TENCON*

2004, Volume C, 21-24 Nov. 2004 Page(s):101 - 104 Vol. 3.

- [VH05] Vanderbauwhede, W.A.; Harle, D.A., “Architecture, Design, and Modeling of the OPSnet Asynchronous Optical Packet Switching Node”, *Journal of Lightwave Technology*, Volume 23, Issue 7, July 2005, Page(s):2215 – 2228.
- [VJS02] Vokkarane, V.A.; Jue, J.P.; Sitaraman, S.; “Burst segmentation: an approach for reducing packet loss in optical burst switched networks”, *IEEE International Conference on Communications, 2002. ICC 2002*, Volume: 5, 28 April-2 May 2002 Page(s): 2673 -2677 vol.5.
- [WD96] Wauters, N.; Demeester, P., “Design of the Optical Path Layer in Multiwavelength Cross-Connected Networks”, *IEEE Journal on Selected Areas in Communications*, Volume: 14, Issue: 5, June 1996, Pages: 881 – 892.
- [WM00] Wei, J.Y.; McFarland, R.I., Jr., “Just-in-time signaling for WDM optical burst switching networks”, *Journal of Lightwave Technology*, Volume 18, Issue 12, Dec 2000 Page(s):2019 – 2037.
- [WPT98] Walter Willinger, Vern Paxson, and Murad Taqqu, “Self-similarity and Heavy Tails: Structural Modeling of Network Traffic”, In *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Adler, R., Feldman, R., and Taqqu, M.S., editors, Birkhauser, 1998.
- [XGY01] Rong Xu; Qian Gong; Peida Ye, “A novel IP with MPLS over WDM-based broad-band wavelength switched IP network”, *Journal of Lightwave Technology*, Volume 19, Issue 5, May 2001 Page(s):596 – 602.
- [XQL04] Jinhui Xu; Chunming Qiao; Li, J.; Guang Xu; “Efficient burst scheduling algorithms in optical burst-switched networks using geometric techniques”, *IEEE Journal on Selected Areas in Communications*, Volume 22, Issue 9, Nov. 2004 Page(s):1796 – 1811.
- [XVC00] Yijun Xiong; Vandenhoute, M.; Cankaya, H.C., “Control architecture in optical burst-switched WDM networks”, *IEEE Journal on Selected Areas in Communications*, Volume 18, Issue 10, Oct. 2000 Page(s):1838 – 1851.
- [YCQ02] Xiang Yu; Yang Chen; Chunming Qiao; “Performance evaluation of optical burst switching with assembled burst traffic input”, *IEEE Global Telecommunications Conference, 2002. GLOBECOM '02*. Volume 3, 17-21 Nov. 2002 Page(s):2318 - 2322 vol.3.
- [YLC04] Xiang Yu; Jikai Li; Xiaojun Cao; Yang Chen; Chunming Qiao, “Traffic statistics and performance evaluation in optical burst switched networks”, *Journal of Lightwave Technology*, Volume 22, Issue 12, Dec. 2004 Page(s):2722 – 2738.
- [YLG01] Tze-Wei Yeow; Law, K.L.E.; Goldenberg, A., “MEMS optical switches”, *IEEE Communications Magazine*, Volume 39, Issue 11, Nov. 2001, Page(s):158 – 163.
- [Yoo96] Yoo, S.J.B., “Wavelength conversion technologies for WDM network applications”, *Journal of Lightwave Technology*, Volume 14, Issue 6, June 1996, Page(s):955 – 966.

- [YPS05] Jie Yu; Petropulu, A.A.; Sethu, H.; “Rate-limited EAFRP-a new improved model for high-speed network traffic”, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]Signal Processing, Volume 53, Issue 2, Part 1, Feb. 2005 Page(s):505 – 522.
- [YQ98] M. Yoo and C. Qiao, “A new optical burst switching (OBS) protocol for supporting quality of service,” in SPIE Proceeding of All Optical Communication System Architecture, Control Network Issues, vol. 3531, Nov. 1998, pp.396–405.
- [YQ99] M. Yoo and C. Qiao, “Supporting multiple classes of services in IP over WDM networks”, in Proceeding of IEEE Globecom, 1999, pp. 1023–1027.
- [YQD00] M. Yoo, C. Qiao, and S. Dixit. “QoS performance of optical burst switching in IP-over-WDM networks”. Journal on Selected Areas in Communications, 18(10):2062-2071, October 2000.
- [YXB02] Shun Yao; Fei Xue; Mukherjee, B.; Yoo, S.J.B.; Dixit, S., “Electrical ingress buffering and traffic aggregation for optical packet switching and their effect on TCP-level performance in optical mesh networks”, IEEE Communications Magazine, Volume 40, Issue 9, Sep 2002, Page(s):66 – 72.
- [YYT05] Yano, M.; Yamagishi, F.; Tsuda, T., “Optical MEMS for photonic switching-compact and stable optical crossconnect switches for simple, fast, and flexible wavelength applications in recent photonic networks”, IEEE Journal of Selected Topics in Quantum Electronics, Volume 11, Issue 2, March-April 2005 Page(s):383 – 394.
- [ZB03] Zapata, A.; Bayvel, P., “Impact of burst aggregation schemes on delay in optical burst switched networks”, The 16th Annual Meeting of the IEEE Lasers and Electro-Optics Society, 2003. LEOS 2003. Volume 1, 27-28 Oct. 2003, Page(s):57 - 58 vol.1.
- [ZOM03] Hui Zang; Ou, C.; Mukherjee, B., “Path-protection routing and wavelength assignment (RWA) in WDM mesh networks under duct-layer constraints”, IEEE/ACM Transactions on Networking, Volume 11, Issue 2, April 2003, Page(s):248 – 258.
- [ZRM03] Zhang, Z.-L.; Ribeiro, V.J.; Moon, S.; Diot, C.; “Small-time scaling behaviors of Internet backbone traffic: an empirical study”, IEEE INFOCOM 2003. Volume 3, 30 March-3 April 2003 Page(s): 1826 - 1836 vol.3.
- [ZS99] Sahinoglu, Z.; Tekinay, S.; “On multimedia networks: self-similar traffic and network performance”, IEEE Communications Magazine, Volume 37, Issue 1, Jan. 1999, Page(s):48 – 52.
- [ZTZ05] Guodong Zhang; Stango, J.T.; Xiupu Zhang; Chongjin Xie; “Impact of fiber nonlinearity on PMD penalty in DWDM transmission systems”, IEEE Photonics Technology Letters, Volume 17, Issue 2, Feb. 2005, Page(s):501 – 503.