ENERGY EFFICIENT WIRED NETWORKING

A DOCTOR OF PHILOSOPHY THESIS SUBMITTED TO THE DEPARTMENT OF

ELECTRONIC ENGINEERING AND COMPUTER SCIENCE

QUEEN MARY UNIVERSITY OF LONDON

BY

XIN CHEN

September 2014

I, Xin Chen, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature:

Date:

ABSTRACT

This research proposes a new dynamic energy management framework for a backbone Internet Protocol over Dense Wavelength Division Multiplexing (IP over DWDM) network. Maintaining the logical IP-layer topology is a key constraint of our architecture whilst saving energy by infrastructure sleeping and virtual router migration.

The traffic demand in a Tier 2/3 network typically has a regular diurnal pattern based on people's activities, which is high in working hours and much lighter during hours associated with sleep. When the traffic demand is light, virtual router instances can be consolidated to a smaller set of physical platforms and the unneeded physical platforms can be put to sleep to save energy. As the traffic demand increases the sleeping physical platforms can be re-awoken in order to host virtual router instances and so maintain quality of service.

Since the IP-layer topology remains unchanged throughout virtual router migration in our framework, there is no network disruption or discontinuities when the physical platforms enter or leave hibernation. However, this migration places extra demands on the optical layer as additional connections are needed to preserve the logical IP-layer topology whilst forwarding traffic to the new virtual router location. Consequently, dynamic optical connection management is needed for the new framework.

Two important issues are considered in the framework, i.e. when to trigger the virtual router migration and where to move virtual router instances to? For the first issue, a reactive mechanism is used to trigger the virtual router migration by monitoring the network state. Then, a new evolutionary-based algorithm called VRM_MOEA is proposed for solving the destination physical platform selection problem, which chooses the appropriate location of virtual router instances as traffic demand varies.

A novel hybrid simulation platform is developed to measure the performance of new framework, which is able to capture the functionality of the optical layer, the IP layer data-path and the IP/optical control plane. Simulation results show that the performance of network energy saving depends on many factors, such as network topology, quiet and busy thresholds, and traffic load; however, savings of around 30% are possible with typical medium-sized network topologies.

GLOSSARY OF TERMS

ATM	-	Asynchronous Transfer Mode
BLX	-	Blend Crossover
CAPEX	-	CAPital EXpenditure
CCU	-	Centralized Control Unit
CP	-	Control Plane
CPU	-	Central Processing Unit
DP	-	Data Plane
DPPS	-	Destination Physical Platform Selection
DWDM	-	Dense Wavelength Division Multiplexing
EEE	-	Energy-Efficient Ethernet
FIB	-	Forwarding Information Base
GA	-	Genetic Algorithm
GMLPS	-	Generalized Multi-Protocol Label Switching
IP	-	Internet Protocol
IPTV	-	Internet Protocol TeleVision
ITU	-	International Telecommunication Union
ISP	-	Internet Service Provider
LSA	-	Link State Advertisement
NGN	-	Next Generation Network
NMS	-	Network Management System
NPGA	-	Niched-Pareto Genetic Algorithm
NSFNET	-	National Science Foundation NETwork
NSGA	-	Nondominated Sorting Genetic Algorithm
NSGA- II	-	Nondominated Sorting Genetic Algorithm II
MOEA	-	Multi Objective Evolutionary Algorithm
MAN	-	Metropolitan Area Networks

MOGA	-	Multi-Objective Genetic Algorithm
MOOP	-	Multi-Objective Optimization Problem
O-E-O	-	Optical Electrical Optical
OPEX	-	OPerational EXpenditure
OSPF	-	Open Shortest Path First
P2P	-	Peer-to-Peer
PC	-	Personal Computer
PP	-	Physical Platform
QoS	-	Quality of Service
RIB	-	Routing Information Base
ROADM	-	Reconfigurable Optical Add-Drop Multiplexer
SDH	-	Synchronous Digital Hierarchy
SFE	-	Switch Fabric Elements
SNMP	-	Simple Network Management Protocol
SONET	-	Synchronous Optical NETworking
SPEA	-	Strength Pareto Evolutionary Algorithm
SPEA2	-	Strength Pareto Evolutionary Algorithm 2
SOOP	-	Single Objective Optimization Problem
SPT	-	Shortest Path Tree
VNE	-	Virtual Network Embedding
VoD	-	Video on Demand
VR	-	Virtual Router
VRM	-	Virtual Router Migration
VRM_MOEA	-	Virtual Router Migration Multi-Objective Evolutionary Algorithm
VROOM	-	Virtual ROuter On the Move
WAN	-	Wide Area Networks
WDM	-	Wavelength Division Multiplexing

CONTENTS

Abstract iii
Glossary of Terms v
Contents vii
List of Tables xii
List of Figures xiii
Acknowledgements xvi
CHAPTER 1 INTRODUCTION 1
1.1 Motivation 1
1.2 Research Objectives 3
1.3 Novelty and Contributions 7
1.4 Report Structure 8
1.5 Authorship 9
CHAPTER 2 BACKGROUND AND RELATED WORK 11
2.1 Introduction 11
2.2 A Typical Network Scenario 11
2.2.1 Network Architecture 11
2.2.2 Physical Platform Architecture 15
2.2.3 ROADM Architecture 18
2.3 Related Work 22
2.3.1 Energy Efficient Approaches 22
2.3.1.1 Green Network Devices 23
2.3.1.2 Green Network 27
2.3.2 Infrastructure Sleeping 29
2.3.3 Virtual Router Migration 32
2.3.4 Power Consumption Model 35

2.3.5 Multi-Objective Evolutionary Algorithms 36
2.3.5.1 Multi-Objective Optimization Problem Concepts 37
2.3.5.2 Evolutionary Algorithms 43
2.3.5.3 Key EA Components 45
2.3.5.4 Multi-Objective Evolutionary Algorithms 47
CHAPTER 3 NEW DYNAMIC ENERGY MANAGEMENT FRAMEWORK 50
3.1 Introduction 50
3.2 Requirements 51
3.2.1 Centralized Control Unit 51
3.2.1.1 Network State Information Collection 52
3.2.1.2 Compute the Destination Physical Platform Selection Algorithm 52
3.2.1.3 Optical Resource Test 52
3.2.1.4 Coordinate Virtual Router Migration Event 53
3.2.2 Virtual Router Migration 53
3.2.3 PP Sleep Function 54
3.3 Novel Dynamic Energy Management Framework 55
3.4 Virtual Router Migration 59
3.4.1 Single VRM Procedure 60
3.4.2 Multiple VRM Procedure 61
3.5 Destination Physical Platform Selection 63
3.6 Triggering Virtual Router Migration 64
3.7 Dynamic Optical Connection Management 65
3.8 Network Power Consumption Model 67
3.8.1 Individual Equipment Power Consumption Model 67
3.8.2 Total Power Consumption Model 72
3.8.3 Cumulative Energy Consumption Model and Energy Efficiency 72
3.9 Summary 74
CHAPTER 4 DESTINATION PHYSICAL PLATFORM SELECTION ALGORITHM 75
4.1 Introduction 75

4.2 Problem Description 75
4.3 Why MOEA 78
4.4 Destination Physical Platform Selection Algorithm-VRM_MOEA 80
4.4.1 Parameter Design 81
4.4.2 Chromosome Encoding 82
4.4.3 Viability Test 83
4.4.4 VRM_MOEA Steps 85
Step 1: Initialization 86
Step 2: Evaluation 87
Step 3: Fitness Assignment 88
Step 4: Environmental Selection 90
Step 5: Mating Selection 92
Step 6: Reproduction 92
Step 7: Final Solutions 96
4.5 Summary 96
CHAPTER 5 SIMULATION MODELLING AND PARAMETER SETTING 98
5.1 Introduction 98
5.2 Simulation Tool 98
5.2.1 Simulation Entities 100
5.2.2 Components of Discrete-Event Driven Simulation 100
5.2.3 Simulation Events 102
5.2.4 Event Graph 103
5.2.5 Simulation Flow Chart 104
5.3 Network Topology 106
5.3.1 Reference Network Topology 106
5.3.2 Random Network Topology Generator 106
5.4 Traffic Model 109
5.5 Random Number Generator 111
5.6 VRM_MOEA Parameter Setting 111

5.6.1 Objective and Performance Metrics 112
5.6.2 Parameter Impact 114
5.6.2.1 Mutation and Crossover 114
5.6.2.2 Maximum Generation 119
5.6.2.3 Number of Offspring 122
5.6.2.4 Mating Pool Size 125
5.6.2.5 Traffic Load 127
5.6.2.6 Number of Nodes 128
5.6.3 Section Summary 129
5.7 Validation of Simulation 130
5.8 Summary 130
CHAPTER 6 SIMULATION RESULTS AND DISCUSSION 132
6.1 Introduction 132
6.2 Simulation Parameters and Assumptions 132
6.3 Numerical Simulations 135
6.3.1 Impact of Traffic Load 135
6.3.2 Impact of Busy and Quiet Thresholds 140
6.3.3 Impact of Migration Cost Setting 142
6.4 Summary 146
CHAPTER 7 CONCLUSIONS AND FURTHER WORK 147
7.1 Conclusions 147
7.2 Further Work 150
REFERENCES 153
APPENDIX A: TRAFFIC ANANLYSIS 176
Part 1 Motivation 176
Part 2 Related Work 177
Part 3 Network Traffic Data 179
Part 3.1 Data Description 179
Part 3.2 Traffic Modelling 180

Part 3.3 Discussion and Summary 190
APPENDIX B: SIMULATION MODELLING 193
Part 1 Event Routines 193
Part 2 Samples of Traffic Demand from the Abilene Network 197
Part 3 Random Number Generator 199
Part 4 Network Topology Generator Validation 200
Part 5 Simulation Tool Validation 201
APPENDIX C: SIMULATION RESULTS 205
Part 1 BLX-a Results 205
Part 2 Migration Cost Setting 206

LIST OF TABLES

Table 1. Simulation Entities 100
Table 2. Types of Events 103
Table 3. RNTG Algorithm 109
Table 4. The Parameters Associated with Power Consumption Model 134
Table 5. The Parameters Associated with VRM_MOEA 134
Table 6. Configurations 143

LIST OF FIGURES

Figure 1. Three Models 13
Figure 2. Substrate Network and Virtualization Network 14
Figure 3. IP Layer and Optical Layer Model 15
Figure 4. Considered Physical Platform Architecture 16
Figure 5. Functional Representation of a Physical Platform 18
Figure 6. Abstract Representation of an ROADM 19
Figure 7. Conventional ROADM Architecture 20
Figure 8. Illustration of Two Spaces in a General Multi-Objective Optimization Problem 38
Figure 9. Illustration of Pareto Dominance 39
Figure 10. Example Pareto Front in the Objective Space 41
Figure 11. Pareto Front Approximation and Pareto Set Approximation 42
Figure 12. Flowchart of Generalised EA Scheme 45
Figure 13. Transitions between Active and Sleep States 55
Figure 14. Overall Framework Operational Procedure 56
Figure 15. (A) Normal Network Configuration. (B) Energy Efficient Configuration where PP3 is in a Sleep State and VR3 and VR4 are running on PP4. 59
Figure 16. Example Scenario (a) Before and (b) After Virtual Router Migration showing the Need for Additional Optical Connections 66
Figure 17. Relationship between Migration Event Time and the Node Power Consumption after Migration 73
Figure 18. An Example Individual / Chromosome 83
Figure 19. VRM_MOEA Flowchart 86
Figure 20. Population Evolution 86
Figure 21. Example of Strength Values of Candidate Solutions 89
Figure 22. Example of Fitness Values of Solutions 90
Figure 23. Example K-nearest Method 92
Figure 24. Example Simple One Point Crossover 94

Figure 25. Interval of Random Generation for a Gene with Blend Crossover 95
Figure 26. Example Blend Crossover 95
Figure 27. Example Mutation Operator on the 4thGene 96
Figure 28. Simulation Event Graph 104
Figure 29. Simulation Flow Chart 105
Figure 30. 6-node-8-link Network Topology 107
Figure 31. NSFNET Network Topology 107
Figure 32. Network Topology of Core Abilene Network 107
Figure 33. Graphical Representation of *Hypervolume* 114
Figure 34. Computational Time of Various Mutation and Crossover Rates 116
Figure 35. The Accumulated Number of Failures of the Viability Test with Various Mutation and Crossover Rates 116
Figure 36. *Hypervolume* with Various Mutation and Crossover Rates 118
Figure 37. Example Crossover Function 119
Figure 38. Impact of Topology on Computational Time of 10N15L Networks 120
Figure 39. Impact of Average Node Degree on Computational Time of 10-node Networks 120
Figure 40. Impact of Random Number Seeds on *Hypervolume* of 10N20L Networks 121
Figure 41. Impact of Network Topology on *Hypervolume* of 10N20L Networks 122
Figure 42. Impact of Network Topology on Computational Time of 10N20L Networks 123
Figure 43. Impact of Network Degree on Computational Time of 10-node Networks 123
Figure 44. Impact of the Number of Offspring of 10-node Networks 125
Figure 45. Impact of Mating Pool Size of 10N25L Networks 126
Figure 46. Impact of Mating Pool Size on Computational Time of 10-node Networks 126
Figure 47. Impact of Mating Pool Size of 10-node Networks 127
Figure 48. Impact of Traffic Load on Computational Time of 30-node Networks 128
Figure 49. Impact of Number of Nodes on Computational Time 128
Figure 50. Energy Saving per Hour with Different Traffic Load in the 6N8L Network 137
Figure 51. Energy Saving per Hour with Different Traffic Load in the 14N21L Network 137

Figure 52. Total Energy Saving with Different Traffic Load in the 6N8L Network..... 138
Figure 53. Total Energy Saving with Different Traffic Load in the 14N21L Network.. 138
Figure 54. Total Energy Saving per Hour of Three Weeks in the Abilene Network 139
Figure 55. Total Energy Saving of Different Threshold Settings..... 142
Figure 56. Additional Optical Hop in the 6N8L Network with Different Threshold Settings..... 142
Figure 57. Number of Occupied Optical Channels of Different Configurations in the 14N21L Network 145
Figure 58. Number of Occupied Optical Channels of Different Configurations in the Abilene Network..... 145

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere appreciation to my supervisor Dr. Chris Phillips for giving me the great opportunity to carry out my PhD study. In the last four years, I really appreciate his continuous guidance, trust, patience and last but not the least, his transfer of knowledge.

Secondly, I would like to thank my good friends Dr. Xian Zhang, Dr. Min Zhou and Dr. Ling Xu, Sabri E. Zaman, Anma Wahid, Dr. Bo Zhong, Dr. Ammar Lilamwala and Yansha Deng for many valuable discussions, encouragement and generous help.

Finally, my special thanks are given to my parents and my Fiancé, Dr. Jian Yao for their consistent support and understandings.

CHAPTER 1 INTRODUCTION

1.1 **Motivation**

Energy efficiency has become an important topic in the last few years and received considerable attention from telecommunication operators, governments and the general public [1]-[6]. The energy consumption of the Internet accounts for approximately 1% of the world's total electricity usage [7]. It is anticipated that this will increase notably over the next few years [8] [9]. Without the adoption of new energy efficient approaches, energy consumption may become one of the main constraints for further growth of the Internet [10].

The continuously rising trend of energy consumption essentially depends on the increasing Internet traffic volume. It is predicted that by the end of 2015, the annual global IP traffic will pass the zettabyte threshold. Global traffic has increased fourfold over the past 5 years and will increase at a rate of 23% over the next 5 years [8]. This trend is driven by the significant increase in the customer population as well as ongoing development of all forms of the Internet-based services, especially bandwidth-intensive applications (IPTV, P2P, etc.).

The unprecedented growth of the Internet brings new challenges to Internet Service Providers (ISPs) and telecom companies. More capable and power-hungry network equipment is required to support these increasing traffic demands. For example, high-end routers are designed in a multi-shelf architecture to provide more network functionalities in a scalable way [11]. Router capacity grows by a factor of 2.5 every 10 months [12]. Meanwhile, the rate of improvement in energy efficiency of silicon technologies is currently about 1.65 every 18 month, which is lower than the factors for the increasing traffic volume or router capacity [13]. Thus, as the improvement in

silicon technologies are not enough to cope with the network energy efficiency issues, new energy efficient approaches need to be explored.

The main motivations that drive the telecommunication network to become "greener" are as follows:

1. Economical

Due to the increasing energy prices, operational expenditure, particularly in terms of the electricity bill, is a significant factor encouraging ISPs and telecom companies to be greener [14]. For instance, British Telecom spent £274m on the electricity bill in 2013 [15]. Even a small energy saving (e.g. 5%) is expected to lead to a significant £13.7m cost reduction. Thus, green networking can reduce ISPs and telecom companies energy costs.

2. Environmental

Energy consumption is an environmental problem. Generally, the world's energy is mainly supplied by burning fossil fuels, which releases carbon dioxide into the atmosphere. It is well known that carbon dioxide is a major component of "Green House" that is a direct contributor of global warming and climate change. Global warming and climate change can result in many environmental problems [17][18]. Hence, reducing energy consumption and increasing energy efficiency is essential for environmental protection and has received considerable attention [16]. For example, a recent European Union (EU) report indicates that the EU needs to more climate-friendly and less energy-consuming to cut most of its greenhouse gas emissions by 2050 [19].

3. Technical

Because of advances in semiconductor and server technologies, the load density of equipment increases, i.e. equipment with more processing and storage capacity can be packed into the same physical space. Meanwhile, the associated energy efficiency is not improved at the similar rate. As a result, the power/heat density increases significantly [20].

The increasing power/heat density has brought many technical challenges to ISPs, especially in terms of cooling. Generally, the heat generated by the equipment must be removed to maintain the desirable operating temperature. However, with the increasing heat/power density, conventional air-cooling systems are reaching their limits [21]. Therefore, it is important to improve the energy efficiency of equipment in order to lower the power/heat density.

1.2 **Research Objectives**

The main research objective is to make a wired access or carrier network operate in a more energy-efficient manner whilst maintaining the same level of performance. Hence, a novel dynamic energy management framework is proposed employing a combination of infrastructure sleeping and virtual router migration.

Within the telecommunications research area, energy/power management can be divided into two parts: wireless networks and wired (fixed) networks. Since the operations of nodes in many wireless networks are limited by the battery power, considerable effort has been devoted to devising new energy management frameworks. As a result, many energy-efficient frameworks have been identified for wireless ad-hoc sensor networks and wireless sensor networks. There is also significant ongoing research on energy saving within cellular networks including the use of low energy (and limited functionality) relay stations, beam steering and exploiting residential wired access points [22].

Conversely, wired networks are generally assumed to have access to abundant power, which implies there is no need to save energy. In reality, power consumption of wired networks is not efficient and has considerable room for improvement [23]. Moreover, technological advances, particularly in respect to the development of Dense Wavelength Division Multiplexing (DWDM), indicate that the bandwidth of physical devices is no longer a restriction on the capacity of the Internet; instead the achievable energy density

limits its development. Therefore, more attention is being given to the energy management within wired networks recently.

Many energy efficient approaches to wired networking have been proposed [1]-[5]. These approaches can generally be divided into two types: network devices and the network itself. The aim of greening network devices is to design energy efficient network equipment by employing energy efficient hardware technologies or modifying the equipment functions, e.g. introducing a sleep/standby state. Green network approaches aim to reduce the amount of energy required to carry out a given task whilst maintaining the same level of performance, such as green network planning [27][33][34], green routing [35][36] and infrastructure sleeping [28].

The Internet can be generally divided into three parts: Tier-3 access networks, Tier-2 Metropolitan Area Networks (MAN) and Tier-1 Wide Area Networks (WAN). Our primary focus is on Tier-2 core networks, which use an Internet Protocol over Dense Wavelength Division Multiplexing (IP-over-DWDM) architecture due to their prevalence and the transport technologies they employ [24]. In addition, although core networks consume a small fraction of the total energy consumption of telecom networks now, they are gradually dominating the energy consumption league because of the increasing popularity of bandwidth-intensive services [25]. The energy/power density issue is also an important driver to increase the energy efficiency of core networks [26].

The network architecture in our study is effectively two conceptually separated networks, i.e. a substrate network and a virtualization network. The substrate network is composed by a group of nodes, each containing a physical platform (PP) and a Reconfigurable Optical Add-Drop Multiplexers (ROADM). The nodes are interconnected by physical fibre. In each node, PPs provide the hardware support for one or more Virtual Router (VR) instances whilst ROADMs offer optical switching and add/drop functions in the optical layer. Above the substrate network, a virtualization network exists, which is composed of a set of VR instances interconnected via virtual

links. A VR instance is an emulation of a physical router, enabling configuration and monitoring of the routing functionality.

In the current core network architecture, routers consume the largest proportion of the energy consumed over the underlying Synchronous Digital Hierarchy (SDH) / Synchronous Optical NETworking (SONET) and Wavelength Division Multiplexing (WDM) layers [24]. Multi-layer models generally increase the complexity and cost of network management and operations. With increasing requirements of service flexibility, reliability and transmission efficiency, it is desirable to eliminate the intermediate SONET/SDH and Asynchronous Transfer Mode (ATM) layers and run IP directly over DWDM [37], i.e. an IP over DWDM network. As a result, routers are required to perform more IP packet processing and consequently consume more energy. Therefore, it is important to particularly improve the energy efficiency of core routers.

One method to improve the energy efficiency is to put routers and interfaces to sleep during the off-peak hours, referred to as infrastructure sleeping. Infrastructure sleeping exploits the characteristics of network traffic. Typically, the traffic demand in a Tier-2/3 network has a regular diurnal pattern based on people's activities, which is high during working hours and much lighter in hours associated with sleep [40] [41]. Moreover, the network architecture is generally over-provisioned and redundant in order to sustain the peak traffic loads and to provide resiliency and fault tolerance. In this case, the network is usually under-utilized in normal operations. For example, the minimum traffic level is typically about 20% to 30% of the peak traffic load [42]. Therefore, infrastructure sleeping has been regarded as a promising strategy and explored intensively in recent years [23][28]-[32].

Despite the promise of infrastructure sleeping, a crucial problem needs to be noted. When a router or a link is in the sleep state, it loses its ability to exchange routing protocol signalling messages with other routers. Subsequently, the logical IP-layer topology changes with the disappearance/reappearance of router links. In consequence,

this change triggers a series of reconvergence events that can cause network discontinuities and disruption.

In order to avoid this problem, we use Virtual Router Migration (VRM) [42][44] to hide the changes in the underlying layer whilst effectively putting some PPs to sleep. A live migration scheme called Virtual ROuters On the Move (VROOM) [44] allows VR instances to move amongst different PPs without causing network discontinuities and instabilities. Nevertheless, VROOM is proposed to reduce the impact of planned maintenance events rather than for energy saving. Furthermore, the decision concerning when to trigger migration and how to determine the appropriate destination PP(s) are not considered nor is the energy consumption is involved. In our study, we discuss how to combine virtual router migration and infrastructure sleeping approaches for energy saving.

In our framework we use a reactive mechanism to determine the time to trigger migration. A centralized control unit is employed to monitor the condition of the network. When significant changes arise, some VRs may need to be moved to new PP locations. In order to limit instabilities due to short term stochastic variations in the load, two thresholds, *i.e. Busy* and *Quiet*, are defined to distinguish between three operational modes for a PP, *i.e. Quiet, Normal* and *Busy*. The network state is observed periodically (e.g. 15 minutes intervals) in order to determine the appropriate time to trigger VRM.

Furthermore, an evolutionary-based algorithm called Virtual Router Migration Multi-Objective Evolutionary Algorithm (VRM_MOEA) is developed to determine the appropriate destination PP if a threshold crossing arises. The outcome of VRM_MOEA is a group of good solutions generated in a relatively short time (e.g. less than 5 minutes) considering various constraints.

1.3 **Novelty and Contributions**

Our research proposes a novel dynamic energy management framework combining the approaches of equipment sleeping and consolidating the functionality of suitable virtual router instances onto fewer platforms that can accommodate them during quiet periods. The research makes the following unique contributions:

1. A new dynamic energy management framework based on an IP-over-DWDM core network architecture is proposed. The main concept is to combine virtual router migration and infrastructure sleeping together with automatic optical layer management to enable resources to be used in an efficient manner. This combination of techniques enables the network to operate in an energy-efficient manner despite large variations in the offered traffic load.

2. A new automatic optical layer management scheme is also developed. The prerequisite for our research is to save energy whilst the logical IP-layer topology remains unaffected. In order to satisfy the requirement, some new optical connections are needed for forwarding the traffic to the remote virtual router instance responsible for processing the packets. Furthermore, these new optical connections also ensure the logical separation of the different virtual router instances that are co-exist on the same physical platform.

3. This research creates a new evolutionary-based algorithm called VRM_MOEA for destination physical platform selection. Particularly, one key question that needs to be resolved within the framework is where to move the virtual router(s) to, based on changing operational circumstances. The approach needs to consolidating functionality of suitable virtual router instances onto fewer physical platforms during the off-peak hours and distribute co-existing virtual router instances away from busy physical platforms during the peak hours. To address this problem, we propose an evolutionary-based algorithm. In the algorithm, a new individual representation is proposed that maps locations of virtual routers onto physical platforms. In addition, two objective functions

are described to evaluate the power consumption and the virtual router migration cost of candidate solutions.

4. A new hybrid simulation platform is developed that is able to capture the functionality of the optical layer, the IP layer data-path and the IP/optical control plane. The optical layer represents channels via a discrete set of integers with associated state information. Meanwhile, a fluid-flow model represents the IP layer data-path and the IP/optical control plane models the signalling messages using discrete packets. The reason a hybrid environment is required is that it is not viable to model IP data-flows as discrete packet streams of packets on an optical infrastructure over many hours.

1.4 **Report Structure**

The rest of the thesis is organized as follows:

Chapter 2 presents background information and a review of relevant literature. A typical IP-over-DWDM network scenario and network components are described in Section 2.2. Then, state-of-the-art related technologies are introduced in Section 2.3 including energy efficiency approaches such as infrastructure sleeping and virtual router migration. Literature concerning the power consumption model and multi-objective evolutionary algorithms is also considered.

Chapter 3 introduces the proposed novel dynamic energy management framework in detail. Initially, some features are required of the network architecture and its capabilities to enable dynamic energy management operations are described in Section 3.2. Then, the overall energy management procedure is described in Section3.3. The related issues of the framework, such as destination physical platform selection, when to trigger virtual router migration and dynamic optical connection management are described in from Section 3.4 to Section 3.8.

Chapter 4 provides a description of the new evolutionary-based algorithm, called Virtual Router Migration Multi-Objective Evolutionary Algorithm (VRM_MOEA) for

destination physical platform selection. The problem is first described in Section 4.2. Then, the motivations for using a MOEA approach are explained in Section 4.3. Finally, details of VRM_MOEA are presented in Section 4.4.

Chapter 5 presents details of the simulation tool and associated network architecture models. We employ a stochastic and discrete-event simulation for driving state-changes within the system. The different functional modules such as the simulation framework, the network topology and traffic models are described in detail from Section 5.2 to Section 5.5. The parameter setting in VRM_MOEA is described in Section 5.6. Finally, the simulation tool is validated in Section 5.7.

Chapter 6 presents the simulation results along with a critical appraisal of the salient features, which are discussed in order to assess the strengths and limitations of the proposed approach.

Chapter 7 considers further work and provides a concise set of conclusions.

1.5 **Authorship**

1. **Chen, X.**; Phillips, C.; "An Evolutionary Based Dynamic Energy Management Framework for IP-over-DWDM Networks," *Sustainable Computing: Informatics and Systems*, March 2014.
2. **Chen, X.**; Wang, J.; Phillips, C.; Reupp, S.; "Dynamic Energy Management Employing Renewable Energy Sources in IP over DWDM Networks, " in *IEEE Online Conference on Green Communications*, October 2013.
3. **Chen, X.**; Phillips, C.; "An Evolutionary Based Dynamic Energy Management Framework for IP-over-DWDM Networks, " in *Energy Efficiency in Large Scale Distributed Systems Conference*, April 2013.
4. **Chen, X.**; Phillips, C.; "Virtual router migration and infrastructure sleeping for energy management of IP over WDM networks," in *2012 International Conference on Telecommunications and Multimedia*, July 2012.

5. **Chen, X.**; Phillips, C.; "Dynamic Energy Management for IP-over-WDM Networks," in *Postgraduate Symposium on the Convergence of Telecommunications, Networking & Broadcasting,* June 2011.
6. Phillips, C.; Gazo-Cervero, Alfonso; Galan-Jimenez, Jaime; **Chen, X.**; "Pro-active energy management for Wide Area Networks," in *IET International Conference on Communication Technology and Application*, October 2011.
7. Wang, J.; **Chen, X**.; Phillips, C.; Yan, Y.; "Energy Efficiency with QoS Control in Dynamic Optical Networks with SDN Enabled Integrated Control Plane," *Computer Networks*, February 2014.

CHAPTER 2 BACKGROUND AND RELATED WORK

2.1 Introduction

The energy efficiency of wired network has been extensively investigated for several years. In Chapter 2, before we describe a novel dynamic energy management framework, appropriate technological background and related work is introduced. Section 2.2 presents a typical network scenario including network architecture (Section 2.2.1), a physical platform architecture (Section 2.2.2) and a reconfigurable optical add-drop multiplexer infrastructure (Section 2.2.3). Next, related work is considered in Section 2.3, this includes energy efficiency approaches (Section 2.3.1), infrastructure sleeping (Section 2.3.2) and virtual router migration (Section 2.3.3), as well as the power consumption model (Section 2.3.4) and the concepts of multi-objective evolutionary algorithms (Section 2.3.5)

2.2 A Typical Network Scenario

2.2.1 Network Architecture

IP traffic today has become dominant as both traditional and new applications such as voice, video and peer-to-peer (P2P) file sharing are employing IP network infrastructure. It is predicted that by the end of 2015, the annual global IP traffic will exceed the zettabyte threshold. Global traffic has increased fourfold over the past five years and will increase at a rate of 23% in the next 5 years [8]. As such, Internet Service

Providers (ISPs) are becoming dissatisfied with the traditional network architecture. Thus ISPs have started to make a slow migration towards a simpler IP-based architecture to handle this unprecedented growth of IP traffic in a more efficient manner [45].

An Internet Protocol (IP) over Dense Wavelength Division Multiplexing (DWDM) network is a promising IP-based architecture because IP provides ubiquitous inter-network transport and DWDM technology offers high and cheap bandwidth [46]. Therefore, the network environment (topology, traffic and routing) in our study is an IP over DWDM core network.

An IP over DWDM network is able to transport the IP traffic over DWDM network with little complexity between these layers. All information and data, no matter it is voice, data or other types of media, is encapsulated into packets using IP. Meanwhile, many of the service functions are separated from the underlying transport technology. The underlying DWDM transport technology multiplexes several carrier optical signals into one fibre using different wavelengths.

Traditional optical network typically contains a Synchronous Optical NETworking (SONET) / Synchronous Digital Hierarchy (SDH) layer [47] and an Asynchronous Transfer Mode (ATM) layer [48] as shown in Figure 1. The SONET/SDH layer provides high speed transmission by transporting voice and data in containers in an efficient time-division multiplexed manner over optical fibres and the ATM layer offers a flexible bandwidth allocation capability. Multi-layer models generally increase the complexity and cost in network management and operations. With increasing requirements of service flexibility, reliability and transmission efficiency, there has been a desire to eliminate the intermediate SONET/SDH and ATM layers and run IP directly over DWDM [49]. As such, compared to traditional optical network architecture, the IP over DWDM network simplifies the network architecture and hence reduces both network OPerational EXpenditure (OPEX) and CAPital EXpenditure (CAPEX).

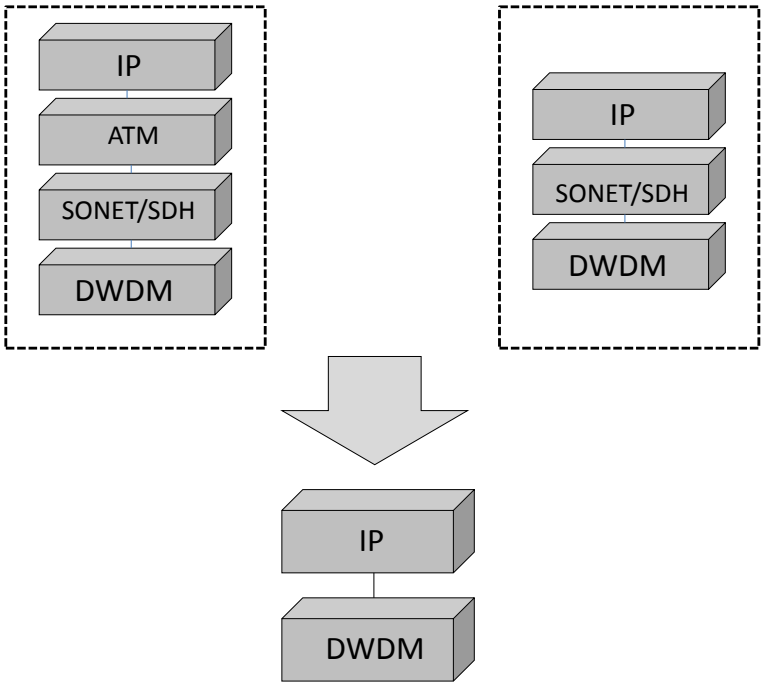

Figure 1. Three Models

A number of researchers extend this concept to effectively regard the architecture as comprising two networks: a substrate network and a virtualized IP network as shown in Figure 2. In our research, we assume that the substrate network is composed of a group of nodes, each node comprising a physical platform (PP) and a Reconfigurable Optical Add-Drop Multiplexers (ROADM). The nodes are interconnected by fibre. Furthermore, in each node, PPs provide the hardware support for one or more virtual router (VR) instances whilst ROADMs offer optical switching and add / drop functions in the optical layer. Meanwhile, the virtualization network exists above the substrate network, which is composed of a set of VR instances interconnected via virtual links. A VR is an emulation of a physical router enabling management, configuration and monitoring of the routing functionality.

As the physical resource of a PP can be shared by several VRs, a substrate network can, in principle, support several virtualization networks. On the other hand, a

virtualization network can also use the resources from several substrate networks. As our study initially explores to combine virtual router migration and infrastructure sleeping for energy saving, for the sake of simplicity, we only consider dynamically distributing a single virtualization network over a substrate network.

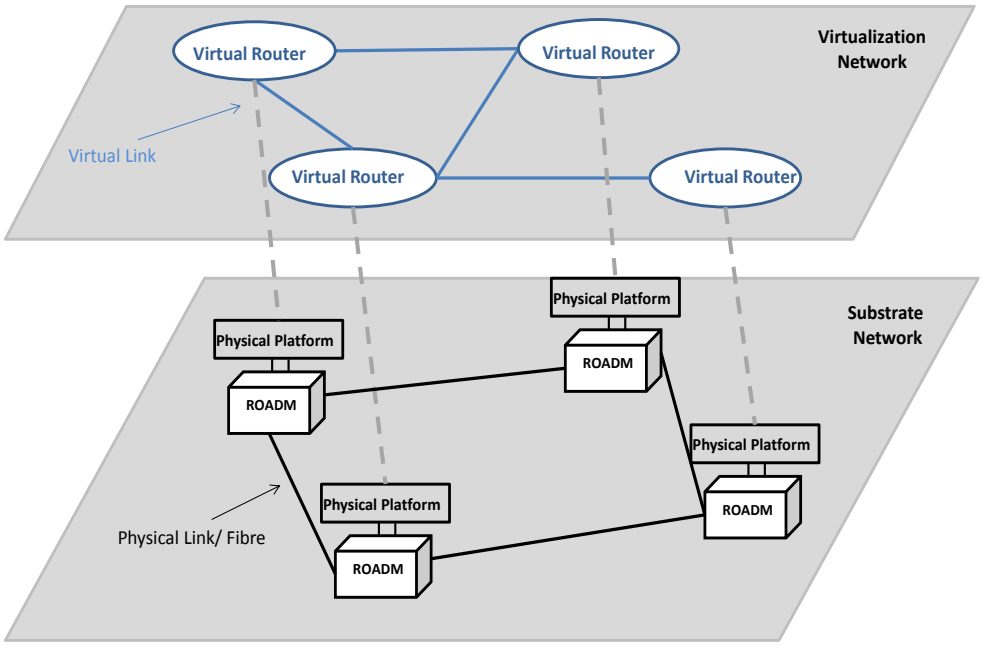

Figure 2. Substrate Network and Virtualization Network

In terms of a layered model, an IP over DWDM network is composed of two principal layers: an IP layer and an optical layer as shown in Figure 3 [16]. In the IP layer, VRs hosted by PPs can process and forward the packets carried by virtual links. The DWDM transponders, which possess Optical-Electrical-Optical (O-E-O) capabilities, can be integrated into PPs [33]. On the other hand, the optical layer provides the inter-connection between PPs via ROADMs. ROADMs are inter-connected with physical fibre links and are responsible for adding and dropping virtual link traffic as well as allowing transit lightpaths/optical connections to bypass PPs in the optical layer. Furthermore, amplifiers can be deployed on the physical fibre links to enable optical signals to transit long distances.

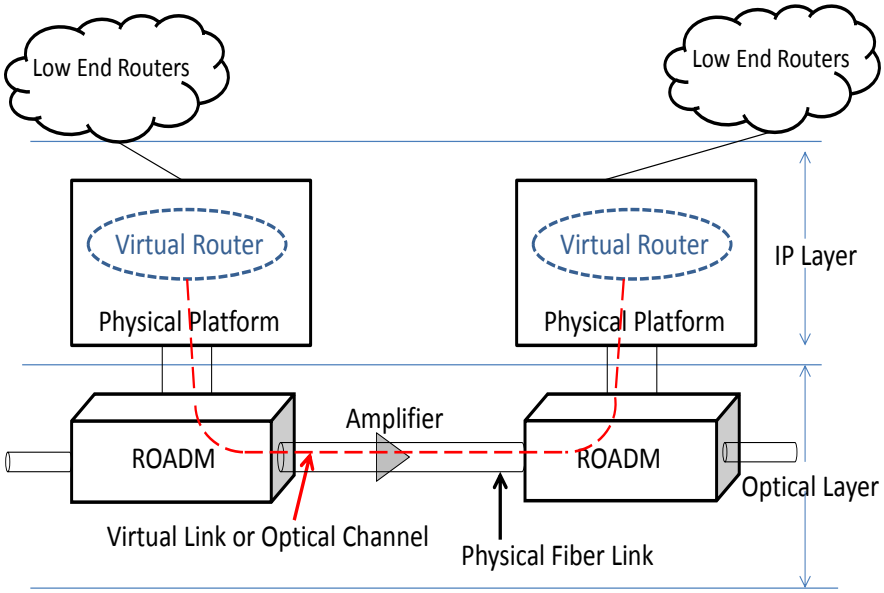

Figure 3. IP Layer and Optical Layer Model

It is important to note that a link in the IP-layer could be composed of several lightpaths in the optical layer. We assume the number of occupied optical channels depends on the actual traffic load dynamically based on an adaptation mechanism proposed in [38]. In the adaptation mechanism, the underlying optical connectivity is adapted by measuring the actual traffic load on lightpaths continuously (periodically based on a measurement period). One or more lightpaths can be established or released depending on the measurement results. Moreover, there is always an optical channel that is reserved for each link to be a control channel for transmitting the signalling messages.

2.2.2 **Physical Platform Architecture**

In our study, a Physical Platform (PP) is a physical router which provides the hardware support for virtual router instance(s). In this section, we use the terms PP and physical router interchangeably.

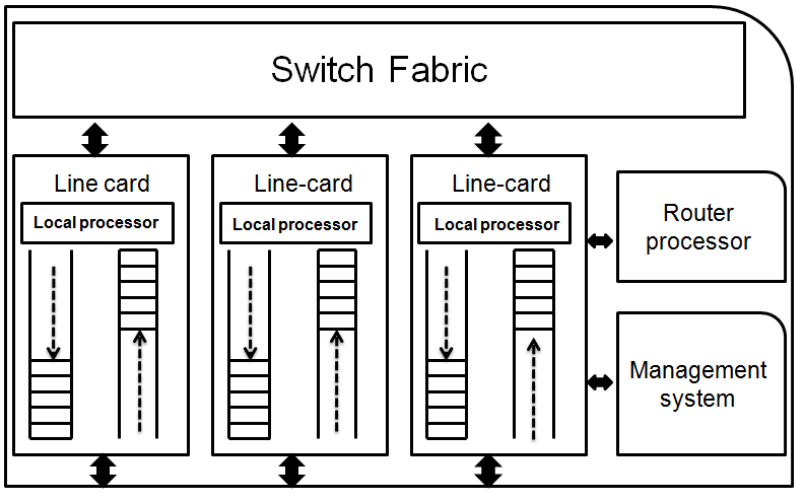

Figure 4. Considered Physical Platform Architecture

A PP is primarily composed of a switch fabric, line-cards, a management system and a router processor (CPU). A typical PP architecture is illustrated in Figure 4. A line-card is responsible for processing and forwarding packets which uses its local processing subsystem (e.g. Forwarding Information Base (FIB)) and buffer space for accommodating the packets via an inbound/ingress interface and an outbound/egress interface. Meanwhile, the switch fabric provides sufficient bandwidth for transferring packets between the different line-cards interfaces. It receives data from line-card ingress interfaces and then switches it to the appropriate line-card egress interfaces. The function of router processor is to maintain the overall forwarding table and distributes relevant parts of the routing table to the line-cards. The management system includes multiple functions, such as cooling, power control and alarm handling.

We assume that a PP has a "sleeping" function which allows the PP to enter a sleep/standby state which consumes low energy. Traditionally, a physical router is designed to remain operational "24/7" and does not possess sleeping capabilities. The trend now is to enable the router or other related equipment to operate with an additional "sleeping" state for energy conservation purposes [28]-[31]. When a router is idle, it can enter the sleep state. Moreover, to be more energy efficient, a line-card is assumed to have a sleep state which consumes energy only when it is working [92].

Of particular importance, we assume that part of the management system always works whether a PP is “awake” or in its sleep state. This part of the management system is responsible for receiving the signalling messages, e.g. enabling a sleeping PP to resume working when the traffic load increases to some pre-defined level requiring it to accommodate a VR instance. Hence, some residual energy is consumed by the management system when a PP is in its sleep state (as noted in the power model in Section 3.8).

In addition, because virtual router migration(VRM) technology is used in our study, three features of the PP architecture need to be satisfied in order to support migration: independence of VR, separation of data plane (DP) and control planes (CP) functionality and the ability to dynamically binding router’s DP to physical substrate interfaces [44].

Independence of VR implies that router instances are decoupled from the physical substrate (i.e. the PP in our framework). The resources of a physical substrate may be segmented across several VR instances. Each VR has its own CP and DP functionality that are isolated from each other. Based on the virtualization feature, the operations of other VR instances co-existing on the same PP are not affected when a VR instance moves to a PP or leaves a PP.

DP and CP functionality separation also means the DP and CP functionality is running in different environment. The DP functionality is to decide how to handle ingress packets by looking up the FIB and then send them to the appropriate egress interfaces. On the other hand, the CP functionality includes generating the network topology, the way to treat packets according to the different service classifications and discarding certain packets. Separation of DP and CP functionality feature already exists in today’s commercial routers. The DP can run on the line-cards and the CP operates on the router processor and main memory.

Dynamic interface binding allows a PP to dynamically establish and change the binding between a VR and physical interfaces. A PP is able to allocate the available interfaces to the VR instance whilst hosting a VR instance. On the other hand, when a

VR instance moves away from a PP, the interfaces associated with the VR instance can be released and may be assigned to other VR instance(s) according to requirements.

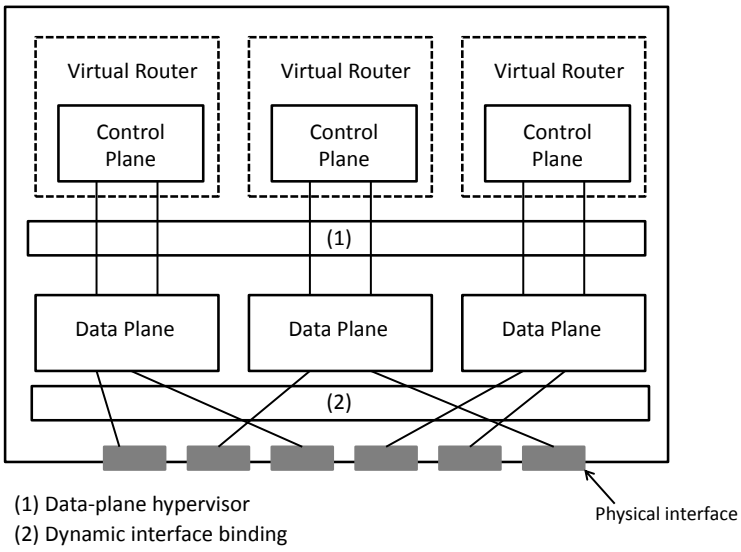

Figure 5. Functional Representation of a Physical Platform

Figure 5 shows a PP from the perspective of its functions including the DP, CP and interface binding modules. Specifically, there is a migration-aware interface called "data-plane hypervisor" proposed in the VROOM system which allows a VR instance to move among several physical routers with different DP technologies.

2.2.3 **ROADM Architecture**

A Reconfigurable Optical Add-Drop Multiplexer (ROADM) is a form of optical add-drop multiplexer which provides flexible bandwidth assignment and configuration of adding, dropping or switching any wavelength to any node at any time when required without affecting the traffic already passing. Meanwhile, an ROADM also offers remote configuration and re-configuration [51]. These functions can be achieved through the use of wavelength blocking, e.g. a Planer Lightwave Circuit (PLC), or Wavelength Selective Switching (WSS). WSS has become the dominant technique because it is cost-effective and scalable [52] [53]. As all the processing of signals is performed in the optical domain eliminating O-E-O conversion, ROADMs induce little delay. ROADMs

have been widely deployed in optical networks to provide automated provisioning which can reduce cost - both OPEX and CAPEX, speed up provisioning time and eliminate manual reconfiguration errors [54].

An example of ROADM functions is shown in Figure 6. A group of wavelengths is processed in the ROADM. The incoming red and purple wavelengths are dropped and new red and purple wavelengths are added in the output interface. The rest of wavelengths pass through the ROADM node. The presence of the ROADM makes the IP over DWDM architecture more flexible as not all packets need to be processed at every node.

ROADMs are often described in terms of degree which usually associated with the number of transmission fibre pairs entering and exiting an ROADM. The degree usually ranges from two to eight. For instance, the directions of a four-degree ROADM are called North (N), East (E), West (W) and South (S). A conventional and typical four-degree ROADM architecture [52] [55] is illustrated in Figure 7.

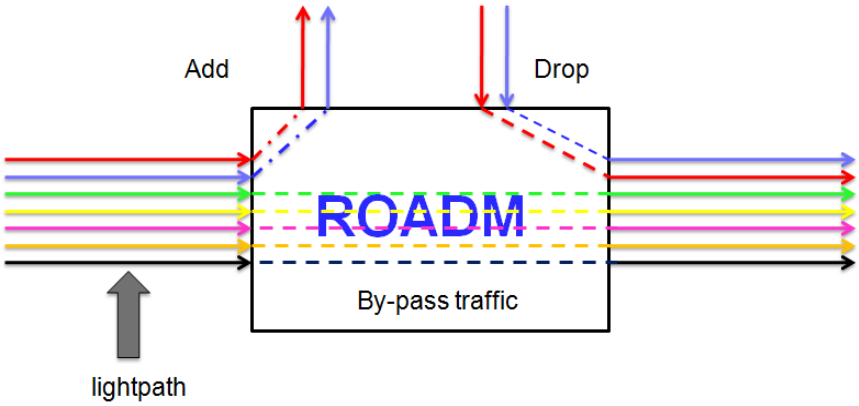

Figure 6. Abstract Representation of an ROADM

There are two main blocks in a conventional ROADM: an optical line switch and an add/drop section. Firstly, the optical line switch is used for routing the incoming wavelengths to their appropriate output ports. In this case, there are four directions on the optical line switch and each direction has a Power Splitter (PS) module and a WSS module. The function of a PS module is to distribute the group of incoming wavelengths

to the WSS modules in the remaining directions as well as the demultiplexer in the drop module for extracting some wavelengths. A WSS module receives groups of wavelengths from PSs in the rest of the directions and wavelengths from the multiplexer in the add module for inserting wavelengths. The WSS selects the wavelengths that should be further transmitted to the network whilst rejecting the remainder. Secondly, the functions of the add/drop section is to insert or extract local wavelengths to/from the network directions. Add/drop functions are achieved by connecting local transponders to the arrayed waveguide gratings (AWGs) that can combine and separate individual wavelength into different physical ports. A transponder is usually full-band tunable and can be provisioned to any wavelength.

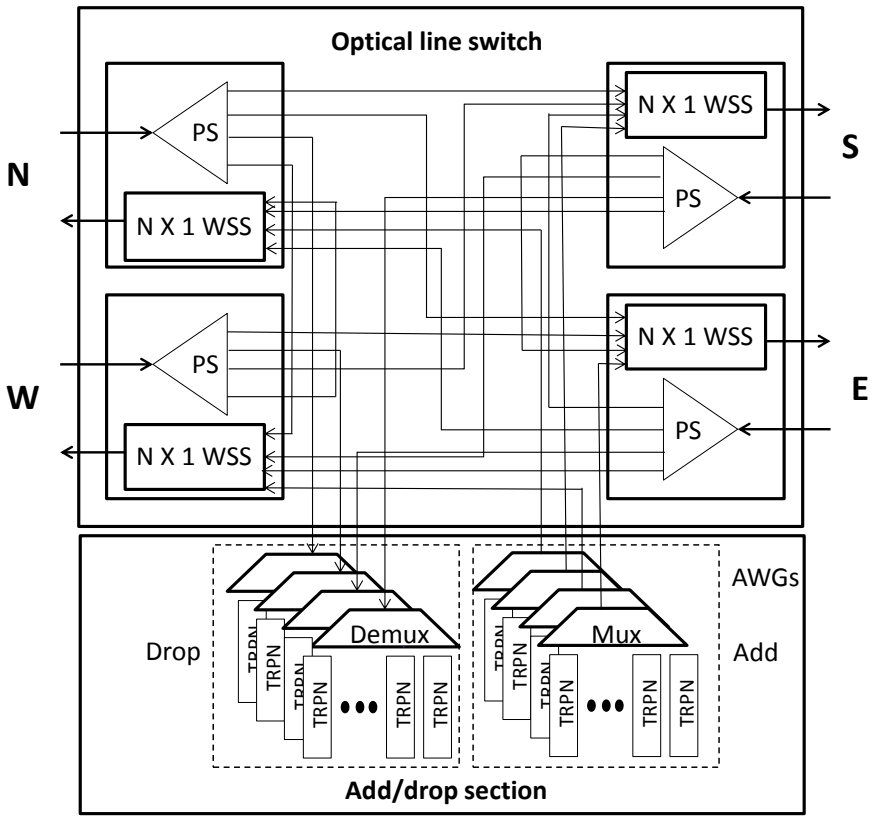

Figure 7. Conventional ROADM Architecture

A conventional ROADM is sometimes referred to as coloured, directional and an ROADM-with-contentions because it is not fully flexible and has three main constraints

in the add/drop section. Firstly, add/drop transponders are required to be assigned fixed wavelengths in order to connect to specific AWG ports, which is known as a "colour" add/drop technology. For example, when a wavelength is selected to be added or dropped, a transponder is required to be adjusted to the wavelength and connected to the correct port of AWGs manually. The second constraint is that an ROADM is directionally dependent. It means that transponders which connect to a specific multiplexer/demultiplexer direction only work for that direction. Hence, it requires a technician to unplug a transponder from one direction to another if the direction of a particular transponder needs to be changed. Finally, because an AWG is not able to process several signals which have the same wavelength, add and drop structures (multiplexer/demultiplexer) are partitioned to avoid the wavelength contention.

Bearing in mind the constraints of a conventional ROADM architecture, extensive studies have been undertaken to enable an ROADM to be completely flexible focusing on architectural implementations and underlying component technologies [55]-[58]. Such an ROADM is called a Colorless-Directionless-Contentionless ROADM (CDC-ROADM). The target of a CDC-ROADM is that any wavelength can be added/ dropped to/from any direction; a transponder can be changed to any wavelength and several signals with same wavelength can be freely added or dropped by any transponder in the add/drop section, without any manual intervention [58]. However, CDC-ROADMs also bring additional node cost and complexity. Although CDC-ROADMs have been discussed for several years, they still need some years to be realized and commercial because they suffers from some problems, such as the power equalization problems due to transients that can arise in a chain of cascaded optical amplifiers.

Despite this, much research in the optical networking is based on a common ideal assumption that an ROADM has sufficient add/drop ports and each port possesses full add/drop capabilities [59]. We therefore assume ROADMs in our study to be CDC-ROADM devices which can provide full flexibility adding, dropping or bypassing any wavelength to any direction at any time.

2.3 Related Work

2.3.1 Energy Efficient Approaches

Energy efficiency is not a new concept. In 1996, a power management standard - Advanced Configuration & Power Interface (ACPI) [60] was published for the Intel 486-DW processor. Since then some energy management mechanisms and hardware enhancements were made but most attention was paid to increasing the efficiency of the Central Processing Unit (CPU).

Around 2003, some pioneering work [23] [61] indicated that the energy efficiency had become an important issue for the further development of the Internet. Compared with the energy efficiency of wireless networks, wired networks have considerable space for improvement [23]. It is estimated that CAPEX and OPEX can be reduced, e.g. by 1 billion dollars every year, if network equipment is "power/energy aware" [61]. There was a relatively "quiet" period on this issue until 2008/2009 when the energy efficiency of the Internet became a hot topic [25] [26] [62] [64]. This was due to the unprecedented growth of the Internet, especially the significant increase in the customer population and the ongoing development of all forms of the Internet-based services. In order to support these increasing traffic demands, ISPs and telecom companies needed to increase network capacity and expand their reach by using more capable and power-hungry network equipment. For example, high-end routers are designed as scalable multi-shelves architecture to provide more network functionalities. Associated with the use of more powerful and power-hungry equipment, the OPEX, particularly in terms of the electricity bill, is a significant factor due to the increasing energy prices. Furthermore, the carbon dioxide emissions, which are associated with electricity generation, cause global warming and have a considerable impact on the environment. As a result, the energy efficiency of the Internet has not only attracted considerable attention from industrial and academic communities, but has also aroused widespread public interest. It is important to note that it is not enough to depend on improvements in

the silicon technology alone to cope with the energy efficiency issues [12][13]. Therefore, in the last few years, various energy efficient approaches have been proposed for wired networks [1]-[3].

These energy efficient approaches can be differentiated into various categories according to different standards. To simplify the taxonomy, we divide these approaches into two types: network devices and schemes that consider the network overall. The aim of network device approaches is to design energy efficient network equipment. On the other hand, overall network approaches focus on network level solutions which can be principally categorized into two subtypes: static and dynamic mechanisms. Static mechanisms consider power/energy consumption an important factor in the network planning and design stage. In contrast, dynamic mechanisms attempt to make better use of the existing network resources by traffic engineering in response to varying traffic loads, such as infrastructure sleeping and power-aware protocols. Note that network device and overall network approaches are not mutually exclusive and they can often be used together to improve the network energy efficiency. As our interest is in IP over DWDM core networks, we mainly focus on the related technologies in this area. The following sections summarize the state-of-the-art of energy efficient network device and network technologies.

2.3.1.1 Green Network Devices

Green network device approaches aim to design energy efficient network elements. This can be achieved by various mechanisms such as:

1. Improve the energy-efficiency of silicon technologies such as Application Specific Integrated Circuits (ASICs), Field Programmable Gate Arrays (FPGAs) and network/packet processors and reduce the internal architectural complexity of equipment [62]-[65];

2. Enable the power consumption of devices to adapt with real traffic conditions or real utilization, such as proportional computing, link rate adaptation and sleep/standby mechanisms [66][67];

3. Replace electrical based devices with pure optical based devices [68] [69].

Some attention has been given to routers to employ more energy efficient hardware technologies and reduce the internal architectural complexity whilst maintaining the same level of performance. One of the promising next generation router architecture is proposed by Roberts [63] suggests that routers should process traffic at the flow level instead of the packet level. If traffic is forwarded at the flow level, a router only need to route the first packet of a flow and the rest of the packets in the same flow can be dispatched quickly. This approach can lead to a simpler and more scalable architecture and thus improve the router energy efficiency. With a similar aim, a synchronous time-based IP switching router is proposed [64], which reduces the router architecture complexity by synchronizing router operations and scheduling traffic demands in advance. The approach has a huge potential to decrease the equipment energy requirements whilst maintaining the same flexibility and performance.

There are several mechanisms which enable the power consumption of a device to adapt to actual network traffic demands or work load such as proportional computing [67], sleep/standby state and link rate adaptation [72].

Proportional Computing – This is when the power/energy consumption of a machine is in proportion to the amount of work performed. For example, a machine consumes nearly no power when it is idle. In addition, the power consumption increases gradually with the increasing workload. Proportional computing can be achieved by tuning the clock frequency (frequency scaling) [70] or/and the voltage of processors (e.g. Dynamic Voltage Scaling) [71]. Frequency scaling and Dynamic Voltage Scaling are able to provide a variable control of energy consumption and the techniques are now widely used in microprocessor design.

Sleep/Standby State – Some sub-components of a device or the entire device can be put in a low power consumption state (i.e. a sleep/standby) rapidly when no activities need to be performed. Conversely, the device resumes working (is re-awoken) when new activities are received. For example, the sleep state is similar to a hibernation mode

of a Personal Computer (PC), which consumes low power when no user applications are running.

In the early position paper [23], a scheme called uncoordinated sleeping is proposed which puts a device into a sleep state based on local information, e.g. during the inter-arrival period between packet transmissions. Two uncoordinated sleeping schemes are discussed: a traditional "wake-on-arrival" approach and a proposed "buffer-and-burst" approach. In the traditional scheme, a sleeping device is wakened when a new packet arrives. Because a transition time is required between the sleep and active modes, energy saving is limited if the packet inter-arrival time is smaller than the transition time. On the other hand, in the "buffer-and-burst" scheme the traffic is shaped into burst-by-burst units and stored in a buffer. The traffic is processed when the buffer is full. This scheme may provide a longer sleeping period whilst the traffic demand is light. The results indicate that the proposed scheme can obtain significant energy savings whilst having little impact on the packet loss and delay. Next, same authors apply the concept on Local Area Network (LAN) switches [73] and Ethernet interfaces [74]. Similarity, an approach is discussed in [75] in terms of grids and servers which improves opportunities to sleep by buffering or (re)scheduling bursts.

Traditionally, network equipment does not possess a sleep function and is designed to keep running "24/7". Due to energy conservation reasons, the trend in the research communities and the industry domain is now to enable routers or other related equipment to operate with an additional sleep/standby state.

Recently, standardization effort has been put into enabling equipment to operate in a more energy-efficient manner. For instance, the IEEE 802.3az Energy-Efficient Ethernet (EEE) standard has been approved by the Institute of Electrical and Electronic Engineers (IEEE) in October, 2010 [76] [77]. EEE provides a hardware support in a physical layer protocol, which reduces the power consumption during times of low data activity. Furthermore, before EEE was approved, an energy efficient feature called Green Ethernet had already been introduced into Ethernet products by some companies [78]. In

addition to the IEEE, the International Telecommunication Union (ITU) has organized a series of symposia on ICT and climate change [79] and the Alliance for Telecommunications Industry Solutions (ATIS) also set up a specific committee to reduce the power consumption of equipment [80].Until now, ATIS has published three standards to determine telecommunication equipment's energy efficiency and suggested a performance measure called the Telecommunications Energy Efficiency Ratio (TEER) [81].

Link Rate Adaptation – The method applies the concept of proportional computing on network interfaces, which allows the capacity of packet processing engines or network interfaces to change dynamically to meet varying traffic demands [72]. Link rate adaptation and sleeping state are compared in terms of QoS such as end-to-end packet delays and losses and energy saving in [28]. The work concludes that both approaches are useful depending on the hardware capabilities of the device and network utilization. For example, an ISP can use link rate adaptation approach during the daytime and infrastructure sleeping at night to satisfy the different traffic requirements.

Another solution for improving the energy efficiency is to replace the current electronic based devices with pure optical based devices, which has been under discussion for many years. Many reviews involve comparisons of energy efficiency among various electronic based and optical based technologies. For instance, several fabric switch technologies of optical and electronic packet switching are compared in [68]. The work summarizes that optical switch fabrics become more energy efficient as the data rate increases. However, the energy efficiency is limited by the optical-electrical-optical (O-E-O) converters required for packet header recognition and header replacement. Similarly, several router architectures are discussed in terms of energy efficiency in [82]. The results indicate that an optical packet switching router which depends on all-optical contention resolution without a conventional store-and-forward mechanism can achieve high energy efficiency. Based on these comparisons, pure optical switching can provide terabits of bandwidth at much lower power dissipation. However, all optical network adoption still requires considerable development. For

example, there are many constraints with pure optical switching architectures such as the limited number of ports and the feasibility of suitable buffering schemes [1] [33].

Although the solution sounds promising, some researchers doubt whether all-optical networking adoption is able to solve the looming "energy bottleneck". For instance, in [10], the benefits of converting buffers and switching fabrics from electronic based to optical ones are discussed. Indeed, photonic buffers and switching fabrics consume much less power than that of electronic ones. However, the power consumption of electronic buffers and switching fabrics only consumes approximately 9% of the entire Internet power consumption. Even if all buffers and switching fabrics were replaced by photonic technologies, the achievable benefit might be less than 9% and hence photonic technologies may not be so desirable.

2.3.1.2 Green Network

In the context of green networks, schemes can be sub-divided into static and dynamic approaches.

1. Static Mechanisms

Static mechanisms take power consumption into account when planning, designing and configuring the network.

One possible approach is to adjust the configuration of chassis and line-cards of core routers appropriately as core routers consume large amounts of energy. An investigation of the power consumption of two existing two generic router platforms is introduced in [33]. Based on measurements, a general router power consumption model is obtained which is composed of the chassis and line-cards. The results indicate that the power consumption varies with different chassis and line-card configurations. Meanwhile, a chassis with a higher fill level of line-cards is more energy efficient. As a result, it is better to reduce the number of chassis that are powered at a given point of presence and to maximize the number of line-cards per chassis from a power-aware perspective.

One paper which takes the lightpath bypass strategy into consideration is [25]. The work uses a Mixed-Integer Linear Programming (MILP) model to minimize the total energy consumption of network components, e.g. IP router ports, transponders and amplifiers. The results indicate that the bypass strategy can save 25% to 45% power consumption compared with the non-bypass design. However, the model is only a rough model focusing on the number of IP router ports used, which degrades the accuracy of the design. A more detailed model of a multi-shelf routing system based on existing commercial network equipment is explored in [27]. The results show that in terms of test networks and traffic matrices, up to 40% energy consumption can be saved.

Aside designing the network infrastructure itself, locating the network equipment more towards strategic places can also reduce energy consumption. Consider a data centre as an example; cooling systems usually contribute a large portion of the overall energy usage. Hence, locating data centres in areas with cold weather or making use of natural resources (water and air) for cooling are prominent ways of cutting power consumption [84]. For example, Google has built their server farms around the Columbia River in the north-western United States in order to make use of local water resources as part of the cooling system [85].

2. Dynamic Mechanisms

Dynamic mechanisms use traffic-engineering methods to make better use of existing resources such as green routing and infrastructure sleeping.

Green Routing – involves designing a new power-aware/energy-aware routing (protocol) which regards energy consumption reduction as an objective. The basic idea is to re-route traffic onto a small set of links during off-peak periods so the remaining links can be powered off. The pioneering work [23] [33] proposes the concepts of power-aware/energy-aware routing which puts components of a router to sleep such as portions of a line-card or an entire line-card during idle times. Later, some heuristic algorithms are proposed to obtain the minimum number of active links [30] [31]. However, the work does not take the routing protocol into account. A green routing

protocol called Energy-Aware Routing (EAR) based on current link-state routing protocols (such as OSPF) is introduced in [35]. EAR suggests only a subset of routers called "Exporter Router" (ER) compute the Shortest Path Tree (SPT) and the remaining routers called "Importer Router" can use ERs' SPT to decide the routing path. Hence, the number of active links is reduced.

In addition, some standardization effort has been put into green routing. New OSPF extensions for MPLS green traffic engineering have been described in a draft for reducing energy consumption [87]. Extensions are used in an Interior Gateway Protocol (IGP)-based central controller framework. The IGP controller collects power consumption information of links and devices and stores it in the IGP extensions. Then, energy efficient paths are computed based on the power consumption information. Note that although green routing allows a network to act in a more energy efficient manner, it requires the support of energy-aware devices (e.g. a sleep function) and network design (e.g. enough path diversity) [86].

Infrastructure Sleeping – These approaches smartly and selectively place some unused network devices in a sleep/standby state and wake them up when necessary. Infrastructure sleeping is an important technology in our study and we therefore describe it in detail in Section 2.3.2.

2.3.2 **Infrastructure Sleeping**

Infrastructure sleeping enables the network equipment to enter a low power consumption state called a standby/sleep state when network activity is low and to re-awaken it, as necessary. In the sleep state, the operational information is retained in order to resume working quickly. Similar to green routing, infrastructure sleeping also requires the support of energy-efficient hardware, e.g. some components or an entire device that has a sleeping capability.

Traditionally, network architecture is designed with over-provisioning and redundancy features to sustain the peak hour traffic load and to provide resiliency and

fault tolerance. In this case, the network is frequently under-utilized in normal operations. Over-provisioned networks offer opportunities for reducing energy consumption. Moreover, due to well-known daily traffic variations [88] [89], typically where traffic varies in an sinusoid-like manner response to people's daily activities, it is possible to put some components or some network devices to sleep during off-peak hours. Thus, infrastructure sleeping appears promising and can achieve significant energy reductions [92] [93].

Despite this, a crucial problem needs to be noted. When a router or a link is in the sleep state (switched off), it loses the ability to exchange routing protocol signalling messages with other routers. This would usually trigger a series of reconvergence events. As a result, these reconvergence events may make the network unstable. Furthermore, the traffic would be re-routed along a longer path, which may be not acceptable due to the congestion, quality of service and the extra latency [90].

As an example, consider the current implementation of Open Shortest Path First (OSPF) protocol [91]. OSPF is a link-state routing protocol and has been widely deployed for many years. The basic mechanisms of OSPF are introduced as follows. By exchanging network state information in the form of Link State Advertisements (LSAs), each router maintains a link state database for computing the local Shortest Path Tree (SPT) based on Shortest Path First (SPF) algorithm. Then, each router makes routing decisions according to its local SPT. Meanwhile, LSAs are exchanged both periodically as well as in response to network state changes in order to maintain the consistency across the network. Hence, if an interface of a router or an entire router enters the sleep state, it cannot exchange messages with its neighbours. Subsequently, its neighbours will generate LSAs to indicate a failed link(s). This is followed by a flood of LSA updates (messages) informing the active routers that the network state has changed. Next, every active router needs to re-computes the SPT. This is a CPU intensive task in a big network which could lead to network discontinuity and disruption. Therefore, how to handle network connectivity is important in infrastructure sleeping approach.

In [23], a coordinated sleeping scheme is introduced which takes a global network-wide energy saving perspective. The basic idea is to re-route traffic via a smaller set of equipment and then some devices can enter a sleep state. In the case, coordinated sleeping collaborates with green routing to avoid the connectivity problem. Later, L. Chiaraviglio *et al*. propose several heuristic algorithms to improve the energy saving by switching off idle nodes and links [29] [30]. The results show that it is possible to switch off 30% of links and 50% of nodes during the off-peak hours. The same authors extend their work in a real-world scenario which employs the actual traffic profiles from the largest ISP in Italy [31]. The results show that an overall energy consumption reduction of around 23% can be achieved, amounting to 3GWh/year. However, as the work uses off-line algorithms with known traffic demand matrices, the connectively problem is not addressed.

Another approach to avoiding the connectivity problem is to use a virtual machine migration which can maintain the IP-layer topology unaffected. For example, some researchers propose to put some line-cards of core routers to sleep during off-peak hours and wake them up as necessary [42]. In their work, a centralized Network Control Unit (NCU) is used to monitor the network conditions. If the traffic is lower than a certain threshold, the NCU activates a sleeping strategy which enables some idle line-cards to go to sleep. In addition, before a line-card enters a sleep state, all functionalities and resources are moved to another line-card within the same physical router to avoid connectivity problems. However, we argue that the base system of a router (a chassis, a router processor and a switching fabric) usually dominates the power consumption of the entire router. Therefore, from energy saving perspective, putting an entire physical router to sleep can conserve more energy.

Based on network redundancy and overprovisioning features, Fisher *et al.* [94] propose a solution which shuts down the redundant cables and line-cards of routers instead of the entire link during periods of low utilization. They indicate that many links in the core networks are actually “bundles” of multiple physical cables and line-cards which can be switched off independently. Besides router cables and line-cards, some

optical devices, e.g. amplifiers and optical switches, are redundant which are installed for protection purposes only. Muhammad *et al*. [97] suggest that these redundant optical devices can be put in a sleep state to reduce energy consumption and they can be promptly reawoken (if necessary) upon a failure.

In addition, the lightpath could be switched off for energy saving in the context of pure optical networks. The work [95] [96] proposes updating the virtual optical layer topology according to traffic demands variation dynamically by adding or deleting lightpaths. The traffic demand of each lightpath is monitored and a threshold-based scheme is applied. If the traffic demand is lower than a threshold, the lightpath can be removed from the virtual optical layer topology. In the opposite case, a new lightpath can be added into the virtual topology when the traffic becomes higher than the threshold.

2.3.3 **Virtual Router Migration**

Another field related to our research is network virtualization, more specifically, Virtual Router Migration (VRM). Since we want to avoid topology change as explained in Section 2.3.2, VRM is used for maintaining the layer-3 IP topology. When the traffic demand is low, Virtual Router instances (VRs) can be moved to a subset of Physical Platforms (PP) and the unused PPs can enter a sleep state to save energy. On the other hand, if the traffic demand increases, sleeping PPs can resume working and VRs are moved to appropriate PPs. In this section, some VRM related technologies are introduced from the concept of virtualization to live virtual router migration systems.

Virtualization is an approach which regroups a set of mechanisms allowing more than one service to operate on the same piece of hardware [2]. Similarly, Network Virtualization(NV) means that multiple virtual networks to co-exist on the same substrate/physical network in order to save space, reduce hardware costs and increase resource utilization [98] [99]. For instance, several ISPs can operate their networks on the top of the same physical infrastructure using different virtual network instances

whilst not interrupting or interfering with each other. In reality, NV has been deployed successfully in the United States Postal Service system [103]. Another NV example is Virtual Private Network (VPN) [104] which allows a private network to go through a public network by establishing virtual channels for traffic isolation.

An important question for NV is how to distribute several virtual network resources effectively on the same substrate network resources, which is called a Virtual Network Embedding (VNE) problem. VNE has been extensively studied for many years [100]-[102]. In addition, NV can bring energy conservation benefits. For instance, in the current machine implementation (e.g. core routers and servers), a machine under high workload consumes less power than several lightly loaded ones. In this context, NV can improve the energy efficiency of devices by sharing the physical resources.

A further benefit of NV technology is that it allows Virtual Machine Migration (VMM), e.g.Virtual Server Migration (VSM) and Virtual Router Migration (VRM) [105] [106]. VMM not only virtualizes the functionalities, but also provides a clean separation between software and hardware (virtualized machines from the underlying infrastructure) which enables a virtual machine to move among different physical hosts. Based on VMM technology, some studies have been done to explore the virtual network migration which is to distribute several virtual networks on the same physical infrastructure dynamically and effectively [107] [108].

In our study, we use VRM technology. In order to have a better understanding of VRM, the concept of a Virtual Router (VR) is introduced firstly. A VR instance is a logical router which separates behavioural functionality from the physical router that hosts the entity. Additionally, a VR includes mechanisms and tools for management, configuration, monitoring and maintenance. In reality, router virtualization is already supported in some commercial routers [112] [113]. Once router virtualization is introduced, it is possible to accommodate multiple VRs within the same physical host if the hardware supports it.

Generally, VRM is used for moving a VR instance amongst several physical hosts without impacting QoS. Hence, VRM is more complex than the router virtualization since it requires solving several problems such as how to minimize the outages during VRM and how to realize the link migration after the virtual router moves to the destination [44].

Some researchers have devoted effort to considering VRM for reducing the impact of planned maintenance since the downtime is a primary concern for ISPs. Typically, there are two types of VRM, i.e. regular (cold) and live (hot) VRM. Regular VRM moves a VR instance after it stops working in the initial (source) physical router. When a VR instance arrives at the new (destination) physical router, the VR instance is required to be re-started to commence work. An example of regular VRM called a Router Farm which is similar to a Server Farm, as proposed in [43]. In a Router Farm, customers are re-homed in the destination physical edge router when the source physical router needs to be upgraded. Because a Router Farm is realized by re-instantiating the destination physical router, it can lead to downtime for both the Date Plane (DP) and Control Plane (CP). On the other hand, in a live VRM, a VR instance continues to work on its initial host whilst all the state information is transmitted to the new physical host. That is to say, a dynamic approach is used for maintaining the packet forwarding whilst migration takes place to minimize the service downtime. For instance, a live VRM framework called Virtual ROuter On the Move (VROOM) is proposed [44]. VROOM allows VRs to move among different physical hosts without causing network discontinuities and instabilities. This migration is typically achieved in the following steps:

1. *Tunnel establish*: A tunnel is established between source and destination physical routers.
2. *Control plane migration*: An image of the VR's CP functionality is created and transmitted through the tunnel to the destination physical router.
3. *Data plane cloning*: The CP functionality on the destination physical router repopulates the DP functionality by using the Forwarding Information Base (FIB) and Access Control Lists (ACLs). After new DP cloning finishes, both old and

new elements of DPrun concurrently on the two physical routers. The remote CP functionality on destination physical router is in charge of routing message exchanges and updating the FIB.

4. *Forwarding links duplication*: Establish outgoing virtual links from the destination physical router to the VR's neighbours whilst all data traffic continues to flow through the source physical router. Then, re-direct the incoming links from neighbours to destination physical router asynchronously. During the asynchronous link migration, data traffic starts to flow through the destination physical router whilst the remaining traffic still flows through the source physical router.

5. *Old data plane and forwarding links elimination*: once the traffic flows on the old forwarding links finish, as determined by a timer, the old DP functionality of VR on the old physical router, old forwarding links and established tunnel in Step 1 are removed.

However, because the purpose of VROOM is to reduce the impact of planned network maintenance, it focuses on minimizing the service down time during VRM and it does not include the mechanisms for energy saving, e.g. the events to trigger VRM and an algorithm to determine the appropriate destination physical router(s). That is to say, the issues of when to move the VRs and where to move the VRs to are not discussed in the paper. Therefore, although VROOM provides an important basic proof-of-concept technology for our research, it lacks a number of key features for energy saving purposes.

2.3.4 **Power Consumption Model**

The power consumption model is important in energy efficient studies because it is a key input to measure the performance of the system. In this section, some power consumption models are reviewed. Generally, there are two types of power consumption model, i.e. component based power models and analytical power models.

Component based power models – These models use the power consumption of individual component/equipment to obtain the total power consumption by counting the occurrences of components/equipment [109]. These models have been widely employed in the research to evaluate specific energy efficiency approaches. An early position paper provides an investigation of the power consumption of two Cisco routers with different line card configurations and indicates that the power consumption of a router is composed of the chassis and some installed active line cards [33]. Based on the power consumption model, the proposed solution is to minimize the number of chassis that are powered and to maximize the number of line-cards per chassis. In the context of the optical network, Shen *et al.* [25] propose a model which considers IP router ports, transponders and optical amplifiers. They argue that the processing of an IP router port consumes high power due to O-E-O conversions, thus an optical bypass strategy is preferred in the network design that does not require the router port processing. Moreover, [27] proposes a model which is based on existing commercial multi-shelf CISCO CRS series routers.

Analytical power models – These models estimate the total network power consumption by considering various factors, such as the average hop count, the energy efficiency of the equipment involved, traffic protection and cooling systems. For example, a traffic protection factor equal to 2 is applied on the power consumption of individual equipment when 1+1 protection is considered. The models are described in more detail in the studies [109]-[111].

2.3.5 **Multi-Objective Evolutionary Algorithms**

A Multi-Objective Optimization Problem (MOOP) is common in the real world. For example, consider a production planning case; several objectives are typically required to be satisfied: (1) maximize the production rate;(2) maximize the machine utilization balance; (3) minimize the throughput time and (4) minimize the overall production time. If there is a solution that can satisfy all the objectives concurrently, the problem can be

transformed to be a single objective problem. However, since some objectives typically conflict, it is not possible to optimize all the objectives at the same time. We usually have to find a "trade – off" or compromise.

An Evolutionary Algorithm (EA) is a popular way to solve MOOPs with conflicting objectives by evolving a group of solutions to approximate an optimal set in a single run. In the following sections, the basic principles of a MOOP and EA as well as EA key components are introduced from Section 2.3.5.1 to Section 2.3.5.3. Finally, a brief introduction to MOEA and its development are described.

2.3.5.1 Multi-Objective Optimization Problem Concepts

A Multi-Objective Optimization Problem (MOOP) is an optimization problem involving several objective functions. Mathematically, objective functions can be transformed into a minimization or a maximization problem, e.g. minimize an objective function is equivalent to maximize its negative or vice-versa. Without loss of generality, a MOOP can be formulated in a minimization context as follows [118]:

Definition 1 Multi-objective optimization problem:

$$\min \ \boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = \big(f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), f_3(\boldsymbol{x}), \dots, f_m(\boldsymbol{x})\big) f_i : \mathbb{R}^n \Rightarrow \mathbb{R} \quad for \ \ i = 1,2, \dots, m$$

$$subject \qquad \boldsymbol{h} = \boldsymbol{h}(\boldsymbol{x}) = \big(h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), h_3(\boldsymbol{x}), \dots, h_k(\boldsymbol{x})\big) \leq 0 \tag{2.7}$$

$$where \boldsymbol{x} = (x_1, x_2, x_3, \dots, x_n) \in \boldsymbol{X} \subseteq \mathbb{R}^n$$

$$\boldsymbol{y} = (y_1, y_2, y_3, \dots, y_m) \in \boldsymbol{Y} \subseteq \mathbb{R}^m$$

A MOOP generally includes a set of *decision vectors* or *solutions*, $\boldsymbol{x}$, which is composed of n decision variables $(x_1, x_2, x_3, \dots, x_n)$, a set of m objective functions $\boldsymbol{f}(\boldsymbol{x}) = \big(f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), f_3(\boldsymbol{x}), \dots, f_m(\boldsymbol{x})\big)$, a set of objective vectors $\boldsymbol{y}$ that represents m objective variables $(y_1, y_2, y_3, \dots, y_m)$ and a set of k constraints $\boldsymbol{h}(\boldsymbol{x})$. $\boldsymbol{X}$ is an n-dimensional *decision space* and $\boldsymbol{Y}$ is the m-dimensional *objective space*. $\boldsymbol{X_f}$ denotes a feasible set of decision vectors which is typically defined by constraint functions and $\boldsymbol{Y_f}$

represents the corresponding feasible set of objective vectors. Figure 8 shows the relationship between the two spaces in a general MOOP.

It is easy to compare two solutions in a Single Objective Optimization Problem (SOOP). If a solution has a better performance with regard to the objective function than the other solution, it is superior. The situation to compare two solutions in a MOOP is complex since there are many objectives that need to be considered. In order to compare different solutions, the concept of Pareto dominance is introduced. Pareto dominance is named after Vilfredo Pareto, an Italian economist who proposed the concept in his studies of economics.

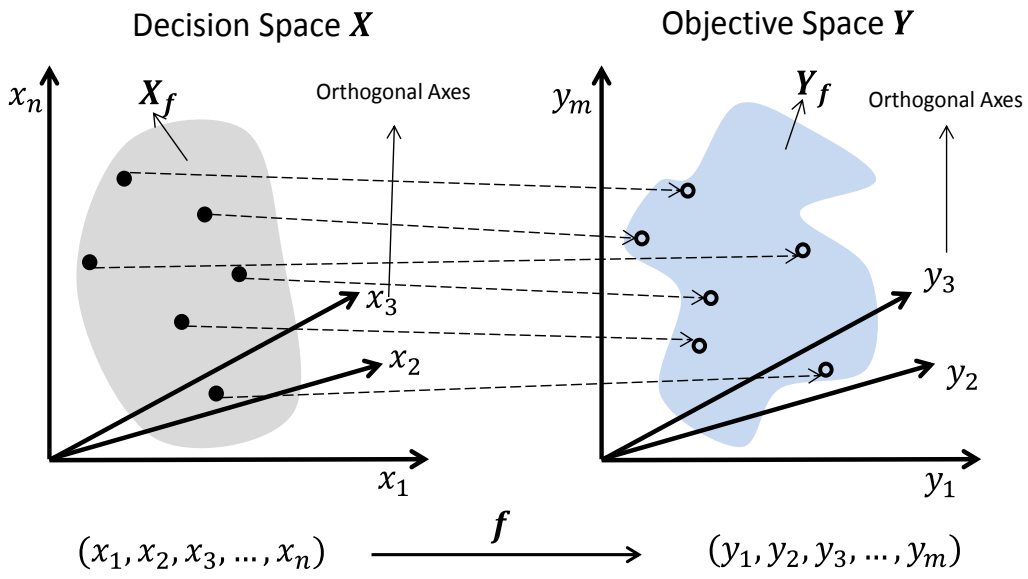

Figure 8. Illustration of Two Spaces in a General Multi-Objective Optimization Problem

For any two decision vectors $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ in a minimization problem, $\boldsymbol{x_1}$ dominates $\boldsymbol{x_2}$ if all components of the objective vector $\boldsymbol{y_1} = \boldsymbol{f}(\boldsymbol{x_1})$ are smaller than the corresponding components of $\boldsymbol{y_2} = \boldsymbol{f}(\boldsymbol{x_2})$. Similarly, $\boldsymbol{x_1}$ weakly dominates $\boldsymbol{x_2}$ if $\boldsymbol{y_1}$ is smaller than $\boldsymbol{y_2}$ for at least one component and not worse in the rest of corresponding components or if $\boldsymbol{y_1}$ equals $\boldsymbol{y_2}$. $\boldsymbol{x_1}$ is indifferent or incomparable to $\boldsymbol{x_2}$ if neither $\boldsymbol{x_1}$ or $\boldsymbol{x_2}$ dominates the other, and they are not equal to one another. That is to say, a / some components of $\boldsymbol{y_1}$ is

smaller than the corresponding components of $\boldsymbol{y_2}$ whilst a / some components of $\boldsymbol{y_2}$ is smaller than the corresponding component(s) of $\boldsymbol{y_1}$. The concept of Pareto dominance in a minimization problem is donated as follows:

Definition 2 Pareto dominance:

$$\begin{aligned} &\boldsymbol{x_1} \succ \boldsymbol{x_2}(\boldsymbol{x_1}\ dominates\ \boldsymbol{x_2}) && iff\boldsymbol{f}(\boldsymbol{x_1}) < \boldsymbol{f}(\boldsymbol{x_2}) \\ &\boldsymbol{x}_{\boldsymbol{1}} \succcurlyeq \boldsymbol{x_2}(\boldsymbol{x_1}\ weakly\ dominates\ \boldsymbol{x_2}) && iff\boldsymbol{f}(\boldsymbol{x_1}) \leq \boldsymbol{f}(\boldsymbol{x_2}) \\ &\boldsymbol{x_1} \sim \boldsymbol{x_2}(\boldsymbol{x_1}\ is\ indifferent\ to\ \boldsymbol{x_2}) && iff\boldsymbol{f}(\boldsymbol{x_1}) \nleq \boldsymbol{f}(\boldsymbol{x_2}) \wedge \boldsymbol{f}(\boldsymbol{x_2}) \nleq \boldsymbol{f}(\boldsymbol{x_1}) \end{aligned} \tag{2.8}$$

An example of Pareto dominance relationship is shown in Figure 9. Assume that two objective functions f_1 and f_2 are both required to be maximized and the points from $\boldsymbol{a}$ to $\boldsymbol{i}$ represent objective vectors in the objective space. The grey area stands for the feasible set $\boldsymbol{Y_f}$ of objective vectors and it can be divided into 4 areas according to the value of two objective functions of $\boldsymbol{a}$.

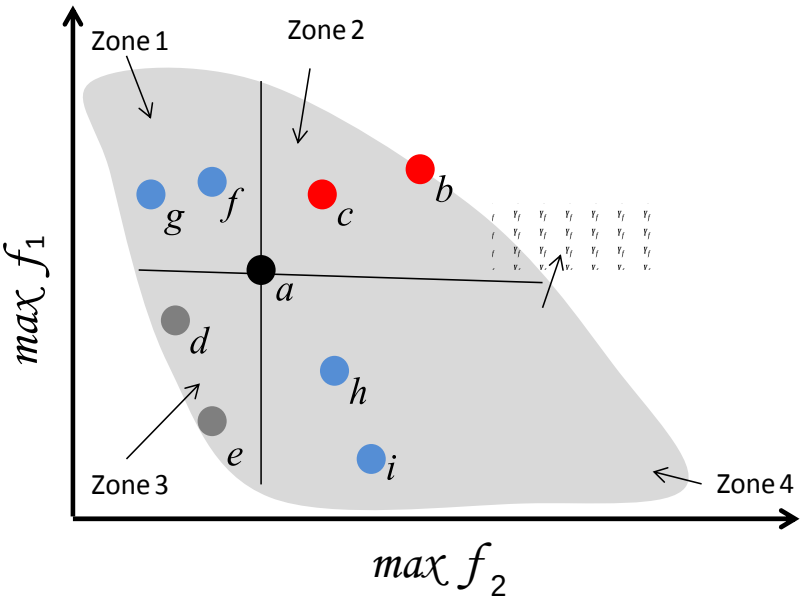

Figure 9. Illustration of Pareto Dominance

Since the values of objective functions of $\boldsymbol{b}$ and $\boldsymbol{c}$ are larger than that of $\boldsymbol{a}$, the corresponding decision vectors of $\boldsymbol{b}$ and $\boldsymbol{c}$ dominate the corresponding decision vector

of $\boldsymbol{a}$. We can deduce that any decision vector whose corresponding objective vector is in *Zone 2* dominates the corresponding decision vector of $\boldsymbol{a}$. Similarly, $\boldsymbol{a}$ has better performance on both objective functions compared with objective vectors in *Zone 3,* so the decision vector represented by $\boldsymbol{a}$ dominates the decision vectors whose corresponding objective vector are in *Zone 3*. For points in *Zone 1* and *Zone 4*, since the value of one objective function is better but the other is worse, decision vectors whose corresponding objective vectors are in *Zone 1* and *Zone 4* are indifferent to the decision vector represented by $\boldsymbol{a}$.

Based on the concepts of Pareto dominance, the optimality criterion of a MOOP can be described. Take Figure 9 as an example, the corresponding decision vector of $\boldsymbol{b}$ is optimal amongst all the points since no other decision vectors can improve an objective value without sacrificing the other objective value. Such a decision vector is called nondominated.

Definition 3 Pareto optimality:

$$A decision\ vector \boldsymbol{x} \in \boldsymbol{X_f}\ is\ called\ nondominated\ in\ a\ set\ \boldsymbol{S} \subseteq \boldsymbol{X_f}\ iff$$

$$\nexists \boldsymbol{s} \in \boldsymbol{S} : \boldsymbol{s} \succ \boldsymbol{x} \tag{2.9}$$

$$if\ the\ set\ \boldsymbol{S}\ equals\ to\ \boldsymbol{X_f}, \boldsymbol{x}\ is\ called\ Pareto\ optimal.$$

Note that if all the objective functions get better or worse at the same time, a MOOP typically can be transformed to be a SOOP. For example, we assign weighted parameters to all objectives and then seek a single optimal solution. However, if the objectives are conflicting (most cases in the real world), e.g. several objectives cannot be optimized concurrently, the situation is complex. In this case, we are seeking a set of trade-offs or compromise solutions instead of a single optimal solution. The set is called a nondominated set:

Definition 4 Nondominated set and Pareto-optimal set:

$$A\ nondominated\ set\ \boldsymbol{A}\ in\ a\ set\ \boldsymbol{S} \subseteq \boldsymbol{X_f}:$$

$$\boldsymbol{A} = \{\boldsymbol{a} \in \boldsymbol{A} | \boldsymbol{a}\ is\ nondominated\ regarding\ \boldsymbol{A}\} \tag{2.10}$$

$if\ the\ set\ \boldsymbol{S}\ equals\ to\ \boldsymbol{X_f}, \boldsymbol{A}\ is\ called\ Pareto-optimal\ set.$

Moreover, the nondominated set of decision vectors' corresponding to the set of objective vectors is called a nondominated front $\boldsymbol{Z}$:

Definition 5 Nondominated fronts and the Pareto-optimal fronts:

$A\ nondominated\ front\ \boldsymbol{Z}$:

$$\boldsymbol{Z} = \{\boldsymbol{z} \in \boldsymbol{Z} | \boldsymbol{z} = \boldsymbol{f}(\boldsymbol{a}), \boldsymbol{a} \in \boldsymbol{A}\} \tag{2.11}$$

$similarily, if\ theset\ \boldsymbol{S}\ equals\ to\ \boldsymbol{X_f}, \boldsymbol{Z}\ is\ called\ Pareto-optimal\ front.$

A Pareto-optimal front in different optimization problems with two objective functions is shown in Figure 10.

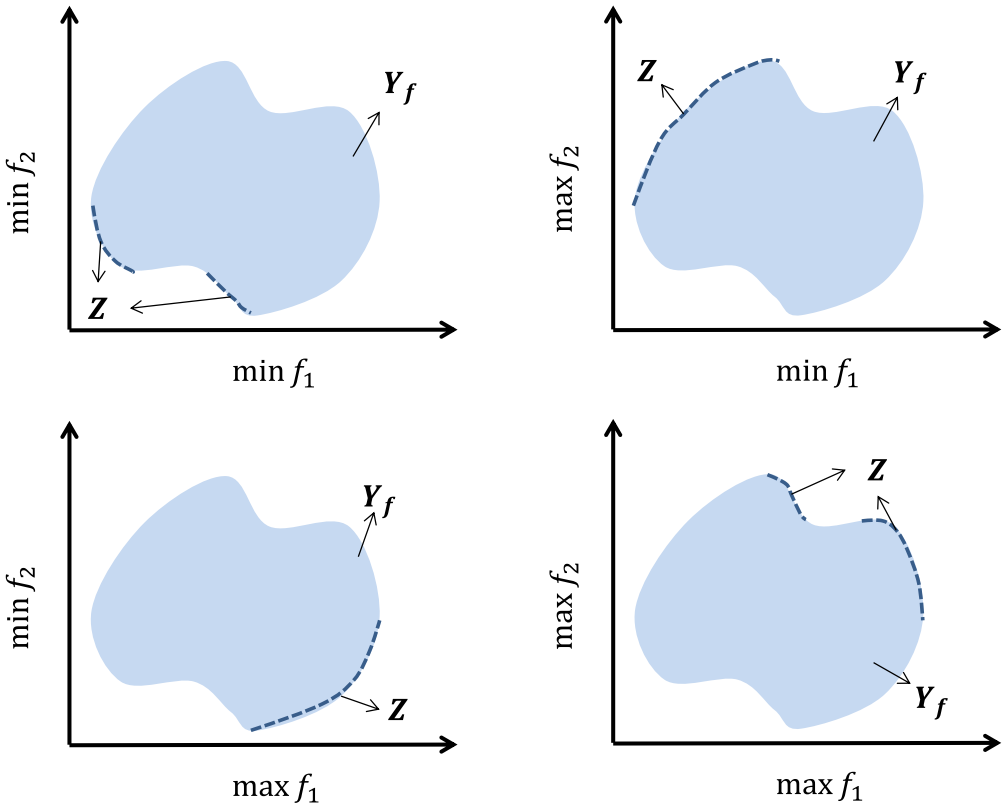

Figure 10. Example Pareto Front in the Objective Space

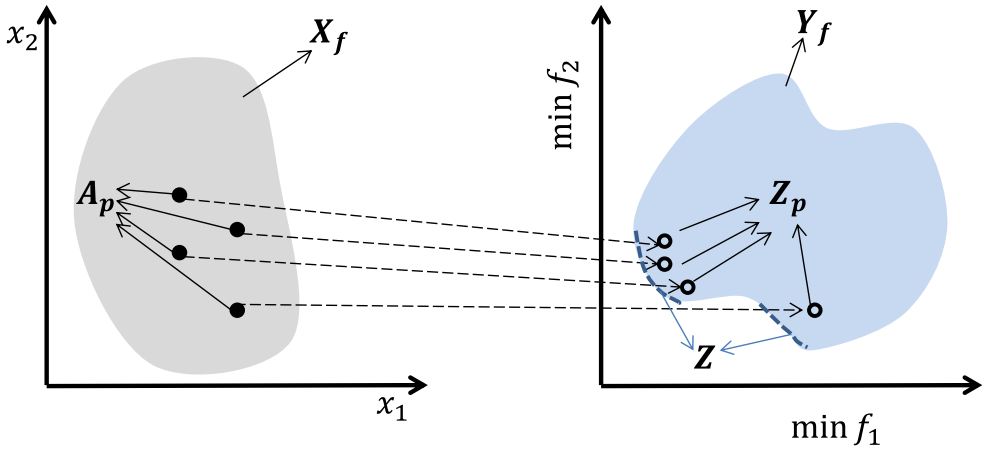

Figure 11. Pareto Front Approximation and Pareto Set Approximation

The goal of a MOOP is to find a true Pareto-optimal set. However, in most problems, it is NP-hard to find the true Pareto-optimal set, which is time consuming and computationally expensive [129]. Thus, the goal of a MOOP is usually to approximate a Pareto set called Pareto set approximation $\boldsymbol{A_p}$. The corresponding set of $\boldsymbol{A_p}$ is called Pareto front approximation $\boldsymbol{Z_p}$. $\boldsymbol{A_p}$ and $\boldsymbol{Z_p}$ are shown in Figure 11.

To solve a MOOP, two sub tasks need be considered: searching and decision-making. Searching is needed to seek the Pareto-optimal solutions. Decision-making involves choosing a suitable final solution from the Pareto-optimal set. Depending on how the search and decision-making processes are combined, three strategies are defined:

- ***Priori:*** The method requires some preference information ahead of the search process. The goal is usually to find a Pareto-optimal solution.
- ***Posterior:*** The method searches (or approximates) the Pareto-optimal solutions without any preference information. Then, a decision maker chooses a suitable final solution.
- ***Interactive:*** The decision maker guides the search process by providing preference information interactively. During each iteration, some trade-offs are presented to the decision maker for selection. After the decision maker provides the preference information, the search continues in the preferred region.

The MOOP approaches we discuss in the report are ***Posterior***, i.e. we find (or approximate) the Pareto-optimal solutions and then perform further operations in the decision making process.

Many classical approaches have been proposed to solve a MOOP, such as the weighting method [115], the constraint method [115] and the goal programming method [116]. Classical approaches usually convert a MOOP to a SOOP. For example, in the weighting method, multiple objective functions are transformed into a linear combination of weighted objective functions. However, classical approaches are generally sensitive to the shape of the Pareto-optimal front or require specific information for adjusting the parameters that may not be available. Moreover, classical approaches usually require several optimization runs to obtain (or approximate) the Pareto-optimal solutions which consumes more computational resources [118] [119].

Employing an Evolutionary Algorithm (EA) is one of the most popular approaches to solving MOOPs because it is able to search an otherwise intractably large and highly complex search space and probably obtain (or approximate) the Pareto-optimal solutions in a single optimization run. EAs are introduced in Section 2.3.5.2.

2.3.5.2 Evolutionary Algorithms

An Evolutionary Algorithm (EA) represents a class of stochastic optimization methods. EA mainly contains several variants such as genetic algorithms (GA), evolutionary programming (EP), Genetic Programming (GP) and Evolution Strategies (ES) [120]. The underlying common idea behind these EA variants is same: use biology-inspired mechanisms such as mutation, crossover and natural selection to refine a set of candidate solutions iteratively.

EAs typically work on a set of candidate solutions called a population. By employing two basic principles: selection and reproduction, the population is modified at each generation. The selection mechanism mimics the fierce competition for survival in the natural world. The fitter individuals have a greater chance of survival and thus to produce offspring. Reproduction includes mutation and crossover genetic operators,

which imitates the process of producing offspring, sharing the genetic information between parent's chromosomes and potentially modifying some genes.

A general EA flow chart is shown in Figure 12. At the start, an initial population is generated by a random or a pre-defined method. Then, candidate solutions in the population are evaluated by the fitness function. Based on their fitness, some candidate solutions are selected to be placed in the mating pool to be parents for further reproduction. The selection process is called mating selection. Note that mating selection is stochastic. This means that fitter candidate solutions have a higher chance of being placed in the mating pool than the less fit ones. However, weak solutions still have a small chance of being selected. When mating selection completes, candidate solutions in the mating pool are used to produce offspring using crossover and mutation operators. When reproduction finishes, a new population (offspring) is generated. Since the offspring are generally created from the fitter candidate solutions, the average quality is normally better than the previous population. Then, in the environmental selection stage, the offspring typically compete with each other and the candidate solutions from the old population to survive through to the next generation. The *selection – reproduction* process repeats until the terminal condition is satisfied. This could be a previously defined computational limit (e.g. a maximum generation count) or when a solution with sufficient quality is found.

EAs benefit from two mechanisms: exploration and exploitation. Exploration is the ability to create population diversity by exploring new solutions in the search space. Conversely, exploitation reduces population diversity by focusing on the candidate solutions with a higher fitness [121]. A common view is that exploration in EAs is performed by genetic operators (mutation, crossover), whilst exploitation is performed by selection [122]. Moreover, it is important to obtain an appropriate balance between exploration and exploitation. Many researchers believe that EAs are successful and effective only when a suitable ratio between exploration and exploitation is determined [123].

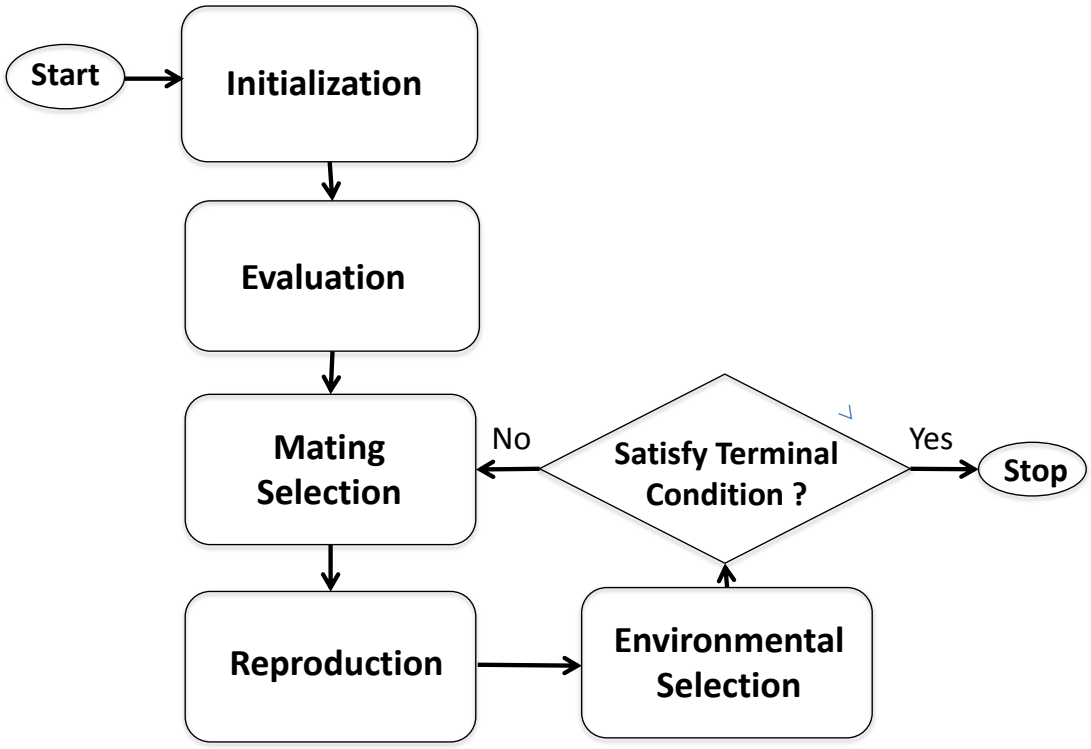

Figure 12. Flowchart of Generalised EA Scheme

EAs have many components and in the following Section 2.3.5.3, some important components are introduced.

2.3.5.3 Key EA Components

Some important EA components are introduced in this section.

1. Individual Representation

The first step in EAs is to link the original problem context to the problem solving space where evolution will take place. The link is achieved by designing an individual representation. A candidate solution in an EA is called an individual or a chromosome. A placeholder in an individual is called a variable, a locus (plural: loci) or gene. The value on such a place is called a value or an allele. The possible solutions in the original problem context are referred to as phenotypes. Mappings of phenotypes, onto individuals in the EAs, are called genotypes. The procedure to design a mapping

between phenotypes and genotypes is called representation. Individual representation is problem specific and encodings can vary, for example, between binary, float and integer.

For instance, in the Travelling Salesman Problem (TSP), each city could be represented by an integer and a candidate solution is a set of integers that stands for the possible route of the salesman's visits [124]. Then one can decide to use a binary encoding to represent an individual, hence city 6 can be seen as a phenotype and "0110" is a genotype.

2. Evaluation Function

The evaluation function or fitness function is used for measuring the quality of individuals. The evaluation function varies depending on the problem context. It is important to note that the concept of an objective function in optimization problems is different from that of a fitness function. An objective function could be identical to an evaluation function if the problem requires maximization because fitness is usually associated with maximization. On the other hand, if the objective function requires minimization, it needs a fitness assignment process for the transformation between the objective function and the fitness function.

3. Mating Selection

Mating selection, or parent selection, aims to select individuals from the current generation to be parents for further reproduction. The structure holding the selected individuals is called a mating pool. Mating selection is typically probabilistic. A fitter individual has a higher chance of being selected than a less fit one but an individual with poor quality still has a small chance of being selected. Some common used mating selection schemes are proportional selection (roulette wheel selection), tournament selection and rank-based selection [125].

4. Reproduction Operator

The role of reproduction is to generate new individuals. Reproduction includes two operators: mutation and crossover (recombination). The crossover operator produces

offspring by exchanging gene information from two parents. The benefit of crossover is that it permits the combination of the best part of each parent to generate offspring. Crossover is stochastic in regard to what parts of the parents' chromosomes are combined and the way these parts are combined typically rely on a series of random choices. The mutation operator maintains the population's genetic diversity by altering some gene values randomly. Similar to the crossover operator, the mutation operator is stochastic. The frequency with which these two operators are invoked is controlled by crossover and mutation rates.

5. Environmental Selection

Environmental selection or the survivor selection mechanism is used to select individuals based their quality (fitness) which is similar to mating selection. As the size of a population is generally a constant, environmental selection is used to decide which individuals can survive through to the next generation depending on their fitness. Environmental selection is often deterministic. For instance, the current population and its offspring are merged and the top segment is selected according to their fitness.

After the introduction of several EA components, some principles and the development of multi-objective evolutionary algorithms are presented inSection 2.3.5.4.

2.3.5.4 Multi-Objective Evolutionary Algorithms

A Multi-objective Evolutionary Algorithm (MOEA) employs evolutionary methodologies to solve problems involving multiple conflicting objectives. MOEAs have been applied in many real world areas, such as economics and finance [127] and engineering [128].

The aim of MOEAs is to find a Pareto-optimal approximation set in a single simulation run. However, it is difficult to get all Pareto-optimal solutions because it is computationally expensive and sometimes it is even infeasible. For example, if the curve of Pareto-optimal front is continuous, the number of solutions is infinite. Hence, the more realistic objective is to find an approximate Pareto-optimal set which satisfies two sub-objectives: 1) the distance from resulting solutions to the Pareto-optimal front

should be minimized; 2) the resulting solutions should have a "good" distribution (usually uniform). In order to realize the above objectives, some important issues need to be considered such as fitness assignment, diversity preservation and elitism [144].

Firstly, compared with SOOPs, where the objective function is sometimes identical to the fitness function, MOOPs need a fitness assignment process after objective values have been calculated for each individual. There are many strategies that can be used for fitness assignment, such as criterion-based and Pareto-based [145].

Secondly, the diversity preservation issue is to ensure that the resulting solutions have a good distribution. This is usually achieved by incorporating density information in the selection process [146]. For instance, an individual has less chance of being selected if it is a short distance from its neighbours.

Thirdly, Elitism aims to avoid losing good solutions during the optimization procedure due to some random effects. There are two common strategies to realize the elitism. One is to merge the current population and its offspring after reproduction into a "temporary" population and then rank the individuals. The individuals in the top segment of the temporary population survive and become the next generation. The other is to establish a special population called the "archive" to hold the promising individuals [136]. The archive is separated from the optimization engine and is updated at each generation when reproduction completes.

The development of MOEAs has seen a number of developments [132]. Initially, popular methods are converted a MOOP into a SOOP using an evolutionary methodology. These algorithms do not incorporate the concept of Pareto optimality. Later, in the mid-1980s, some algorithms started introducing Pareto optimality into the evolutionary algorithms. These algorithms include the Nondominated Sorting Genetic Algorithm (NSGA) [133], the Niched-Pareto Genetic Algorithm (NPGA) [134] and the Multi-Objective Genetic Algorithm (MOGA) [135].

When elitism became a *de facto* standard mechanism in the late 1990s, the new stage of evolution MOEAs began. The landmark algorithm in the field is generally considered

to be the Strength Pareto Evolutionary Algorithm (SPEA), which introduces an external population called the "elite" archive to retain suitable nondominated solutions. This is an important feature as it guarantees that the final solutions are nondominated with respect to all other solutions across the total evolutionary process rather than the current population. After SPEA, many algorithms incorporating additional archive mechanisms have been proposed. The most representative of them are: Strength Pareto Evolutionary Algorithm 2 (SPEA2) [136], Pareto Archived Evolution Strategy (PAES) [137], Nondominated Sorting Genetic Algorithm II (NSGA – II) [138]. MOEAs, such as NSGA-II and SPEA2, have proven to be well-suited for complex MOOPs with two or three objectives [139].

Current research into MOEAs has taken a number of directions. For example, developing new algorithms to deal with problems comprising "many" objectives [140], modifying existing algorithms or hybrids of existing algorithms[141]. These recent developments of MOEAs are covered more fully in [142]-[143].

CHAPTER 3 NEW DYNAMIC ENERGY MANAGEMENT FRAMEWORK

3.1 **Introduction**

Chapter 3 describes a novel dynamic energy management framework including several key features. The main concept of the proposed framework is to combine Virtual Router Migration (VRM) together with selected sleeping of unneeded Physical Platforms (PPs) alongside automatic optical layer management to enable resources to be used in an efficient manner. The logical IP-layer topology remains unchanged because of the VRM. Hence, network disruption and discontinuities are avoided when PPs enter or leave their hibernation state. The remainder of the chapter is structured as follows. Firstly, Section 3.2 introduces several requirements of the network architecture and functionality, which are essential to operations of dynamic energy management. Next, the overall dynamic energy management procedure is described in Section 3.3. A single Virtual Router Migration (VRM) as well as multiple VRM procedures are then described in Section 3.4. Next, two important issues are considered, i.e. where to move the Virtual Router (VR) instance to and when to trigger VRM; these are discussed in Section 3.5 and Section 3.6, respectively. Section 3.7 describes the dynamic optical connection management. This is followed by the network power consumption model in Section 3.8. Finally, a chapter summary is presented in Section 3.9.

3.2 **Requirements**

Our work is an initial exploration of how to reduce network energy consumption without adversely affecting network performance by infrastructure sleeping and virtual router migration. Hence, some features are required of the network architecture and its capabilities to enable dynamic energy management operations. These features are described as follows.

3.2.1 **Centralized Control Unit**

In the framework, a centralized control mechanism is used. Although all monitoring information needs to be brought to a single management unit which imposes a latency and communication overhead, the benefit is that the global knowledge can be obtained. Global knowledge is important in our framework as a single VR cannot make sensible migration decisions alone. Hence, a Centralized Control Unit (CCU) is proposed to supervise and coordinate various functions of dynamic energy management framework. Generally, Simple Network Management Protocol (SNMP) can be used to obtain the network state information. In our case, the CCU, which can be an adjunct to a Network Management Station (NMS), collects the network state information from VRs and PPs using a combination of SNMP *get* requests and *trap* messages. Through these messages, information concerning PPs and links can be ascertained.

As we described in Section 2.2.1, an optical control channel is always maintained for every link. The signalling messages between the CCU and PPs / VR instances / ROADMs are transmitted on these control channels.

The key functions of the CCU and theirassociated procedures are described from Section 3.2.1.1 to Section 3.2.1.4.

3.2.1.1 Network State Information Collection

In the framework, the CCU needs to know the latest state of the network to determine whether to move VRs or not. Thus, network state information is collected periodically, e.g. at 15 minute intervals [95]. The procedure is described as follows.

1. The CCU sends a "state enquiry" message to PPs to obtain the latest state information, e.g. PP state (sleep or active), PP utilization, VR workload, and availability of line-cards. The message is flooded.
2. When a PP receives a "state enquiry" message, the PP collects the latest related data and places it in a "state" message. Then, the PPs send their "state" messages back to the CCU. If a PP is in its sleep state, part of its management module still runs in order to respond to state enquiry messages. In this case, the data related to VR instance is null in the "state" message.
3. When the CCU receives all the "state" messages, it analyzes the data to determine if the network conditions warrant migration.

3.2.1.2 Compute the Destination Physical Platform Selection Algorithm

If the network conditions satisfy the migration condition, the CCU invokes the destination physical platform selection algorithm to choose the appropriate PPs for VR instances. The algorithm is described in detail in Chapter 4.

3.2.1.3 Optical Resource Test

When a group of solutions are obtained from the destination physical platform selection algorithm, the CCU initiates the optical resource test to examine whether the optical resources support a given candidate solution. In the test, the first-fit algorithm is used for wavelength assignment and the wavelength continuity constraint is also considered. The procedure is described as follows.

1. The CCU sends a message to all ROADMs to obtain optical channel availability information.

2. When an ROADM receives the message, it sends a message with the optical channel availability information back to the CCU.
3. When the CCU receives all messages, it computes the optical test to determine whether the optical resources support the candidate solution.

3.2.1.4 Coordinate Virtual Router Migration Event

When the candidate solution is identified, the CCU starts coordinating the VRM mechanism. This contains three stages as follows.

1. Before the VRM starts, if a/some sleeping PPs are needed to host VR instance(s), the CCU sends a "wake" message to these PPs. When sleeping PPs receive a wake message, they resume working.
2. During the VRM procedure, the CCU sends a message to the VR instances which are required to be moved to new locations. When these VR instances receive the message, they start preparing for the migration, such as creating an image of VR's CP functionality. Then, when CP functionality migration completes, the CCU sends a separate message to each VR instance; the message contains its neighbours' information. Further details are provided in Section3.4.
3. When the VRM completes, the CCU sends a "sleep" message to PPs which do not host any VR instance(s). When these PPs receive the "sleep" message, they enter the sleep state.

3.2.2 Virtual Router Migration

Three features are needed to be satisfied in PP architecture: *router virtualization*, *Control Plane (CP) and Data Plane (DP) functionality separation* and *dynamic interface binding*, to make a VR migratable. These features are available to some extent in commercial routers [44].

Router Virtualization – a router instance is separated from its physical substrate (i.e. the PP in our framework). The resources of a physical substrate can be segmented by

several VR instances. Each VR has its own CP functionality and DP functionality that are isolated from each other. Based on the virtualization feature, when a VR instance moves in or out of a PP, operations of the remaining VR instances co-existing on the same PP are not affected.

CP and DP functionality Separation – the CP functionality and DP functionality of a VR instance run in different environments. For instance, the CP is runs on the CPU whilst the DP runs on the line-cards. This separation allows the CP functionality and DP functionality of a VR instance to be moved separately.

Dynamic Interface Binding – this allows a PP to dynamically change the binding between PP interfaces and a VR. A PP is able to allocate available interfaces to a VR instance whilst the PP hosts it. On the other hand, when a VR instance moves away from a PP, the interfaces associated with the VR instance can be released and may be assigned to other VR instance(s) according to requirements.

3.2.3 **PP Sleep Function**

An entire PP need to have power management primitives at the hardware level which have the ability to enter asleep state with lower power consumption. Additionally, PPs should be able to enter the sleep state and resume working (active state) in a short period of time (e.g. few seconds or less).

The sleep function is not supported in existing commercial routers. However, due to the "green" networking, the trend in the research and industry domains is to enable routers or other network equipment to operate with an additional sleep state. For example, the IEEE standard 802.3az Energy Efficient Ethernet (EEE) was approved in 2010. A transition mechanism between active and sleep states of the EEE scheme is shown in Figure 13. The mechanism contains two periods: time to sleep (T_s) and time to wake (T_w). Typically, the default T_w is aimed to be similar to the delay of a maximum length packet at the target link speed [160]. For instance, values of T_s and T_w for a 10Gb/s interface are 4.48 μs and 2.88 μs, respectively [77]. By analogy, we also apply

the transition mechanism between two states based on the EEE scheme. The overhead is the time to put a PP into a sleep state and the time to wake it up, returning them to the active state.

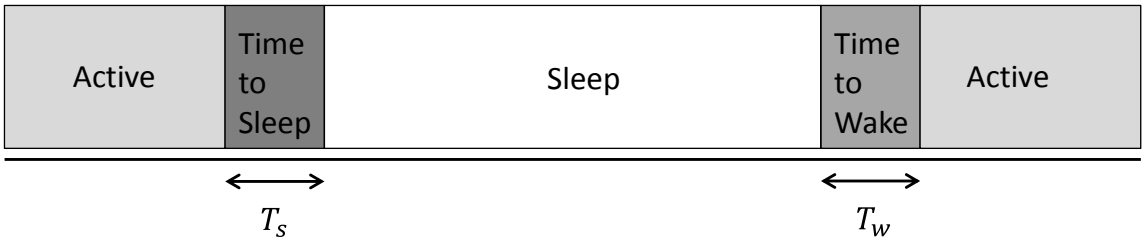

Figure 13. Transitions between Active and Sleep States

3.3 **Novel Dynamic Energy Management Framework**

In this section some terms used in the framework are first defined in order to describe conditions to trigger VRM. Then, the overall dynamic energy management procedure is described.

Firstly, we define that a Physical Platform (PP) has two states: *Active* and *Sleep*. An *Active* state means that components of a PP are working normally. On the other hand, when a PP is in a *Sleep* state, all components are put into hibernation except a small management module, which is used for maintaining and exchanging signalling messages. Next, an active PP has three operating modes: *Quiet*, *Normal* and *Busy* to indicate how busy the machine is. In order to provide hysteresis to prevent unnecessary frequent transitionsdue to short-term stochastic variations in the load, two thresholds, *i.e. Busy* and *Quiet*, are defined. These two thresholds are also used as a condition to trigger VRM. If some PPs are effectively underutilized crossing their *Quiet* threshold, it is preferable to consolidate their VRs on fewer remaining PPs to save energy. In contrast, when a / some working PP(s) cross the *Busy* threshold and enter *Busy* mode, sleeping PP(s) are then selectively reawakened and some VRs are moved away from PPs where

the traffic load is increasing to an unsustainable level. A further situation is when all active PPs are in their *Normal* mode so no VRM is triggered.

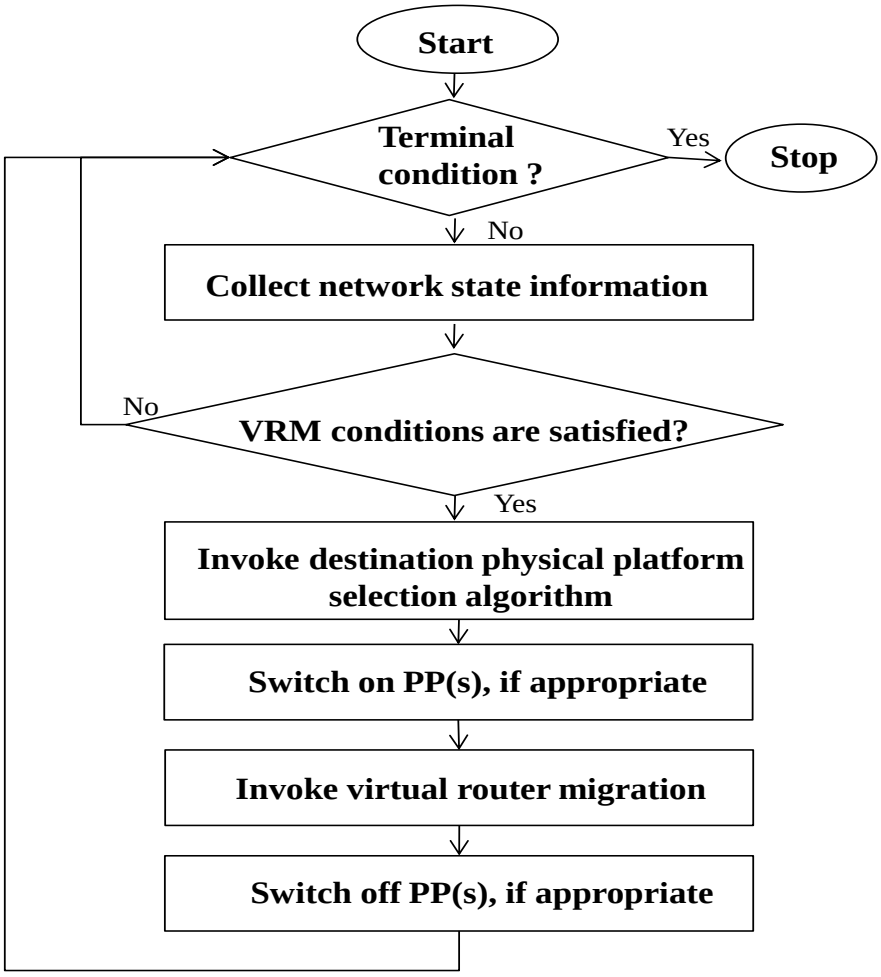

Figure 14. Overall Framework Operational Procedure

Secondly, in order to describe a VR location in the network, three terms: $PP_{default}$, PP_{cur} and PP_{dest} are defined. $PP_{default}$ indicates the original location of a VR and every VR has its own$PP_{default}$. If VRM does not happen, each VR runs on its $PP_{default}$. We assume the traffic that arrives from an access network is always processed by a particular VR (irrespective of the hosting PP). Hence, if a VR moves away from its *default PP*, some additional optical connections are needed to re-direct the traffic to the

remotely located VR instance. A PP_{cur} is where the VR is presently accommodated and a PP_{dest} is the location selected to host the VR after a migration.

The overall operational procedure of the dynamic energy management framework is presented in Figure 14 and the steps are described as follows.

1. **Collect and analyze network conditions**.

A centralized operating model and a reactive mechanismare used in the framework. At first, a centralized operating model is required because an individual VR is not best placed to decide when to move since it does not know the global network conditions. The dynamic energy management framework also needs to coordinate the functioning of PPs and VRs, e.g. sleeping and VRM. Hence, a Centralized Control Unit (CCU, as described in Section 3.2.1) is used for collecting the related network conditions. Typically, the network state is monitored at 15-minute intervals. The information includes VR traffic load, PP utilization and optical resource availability, etc. Secondly, a reactive mechanism is employed to trigger VRM. A reactive mechanism is one where a system performs operations in response to some significant changes in the network. For example, the CCU starts computing the destination PP selection algorithm (as explained in Chapter 4) when it receives a message that a *Busy* or *Quiet* threshold has been crossed.

2. **If the system satisfies migration conditions, then it selects appropriate destination PPs for accommodating VRs.**

The migration conditions include that some PPs are effectively underutilized crossing their *Quiet* threshold or a / some working PP(s) cross the *Busy* threshold and enter *Busy* mode. Under the circumstances, the destination PP selection algorithm is computed. The aim of the algorithm is to maximize the consolidation of VRs onto as few PPs as possible given various constraints. It is necessary to determine which VRs are viable candidates to be migrated so that PP(s) can be placed in a sleep state or to determine when PP(s) need to be re-awoken to accommodate VR(s). The details of destination PP selection algorithm are described in Chapter 4. The output of the algorithm is a solution describing the appropriate locations of VRs.

3. **Switch on PP(s) if appropriate**.

After obtaining a solution from the destination PP selection algorithm, the CCU sends "wake" messages to a / some PP(s) in the sleep state to prepare them for accommodating VR instance(s). It takes a short time (as explained in Section 3.2.3) to wake appropriate PPs up, returning them to the active state.

4. **VR(s) are moved to their appropriate destination(s).**

Based on the identified solution, some VRs are moved to their destination PP locations. VRM procedure is described in detail in Section 3.4.

5. **Switch off PP(s) if appropriate**.

When (multiple) VRM completes, a PP / some PPs do host any VR. Such PP(s) is unneeded PP(s) and can be put into the sleep state to save energy.

6. **Return to step 1 to recheck the network conditions.**

The process repeats indefinitely or within the simulation environment, until a terminal condition is reached. The terminal condition can be predefined simulation duration.

Figure 15 shows two scenarios (a) a normal network configuration and (b) a proposed energy efficient configuration. In this case, 4 VR instances are accommodated by their $PP_{default}$ in scenario (a). Assume that VR3 has a low workload and its $PP_{default}$ PP3 is in *Quiet* mode so the destination PP selection algorithm is computed in order to search for an appropriate location for VR3. The algorithm result shows that VR3 can be hosted by PP4 without detrimentally impacting the performance of VR4. Then, VR3 is moved toPP4. When VRM is complete, both VR3 and VR4 are running on PP4. Finally, PP3 is put into sleep state for energy saving purposes as shown in scenario (b).

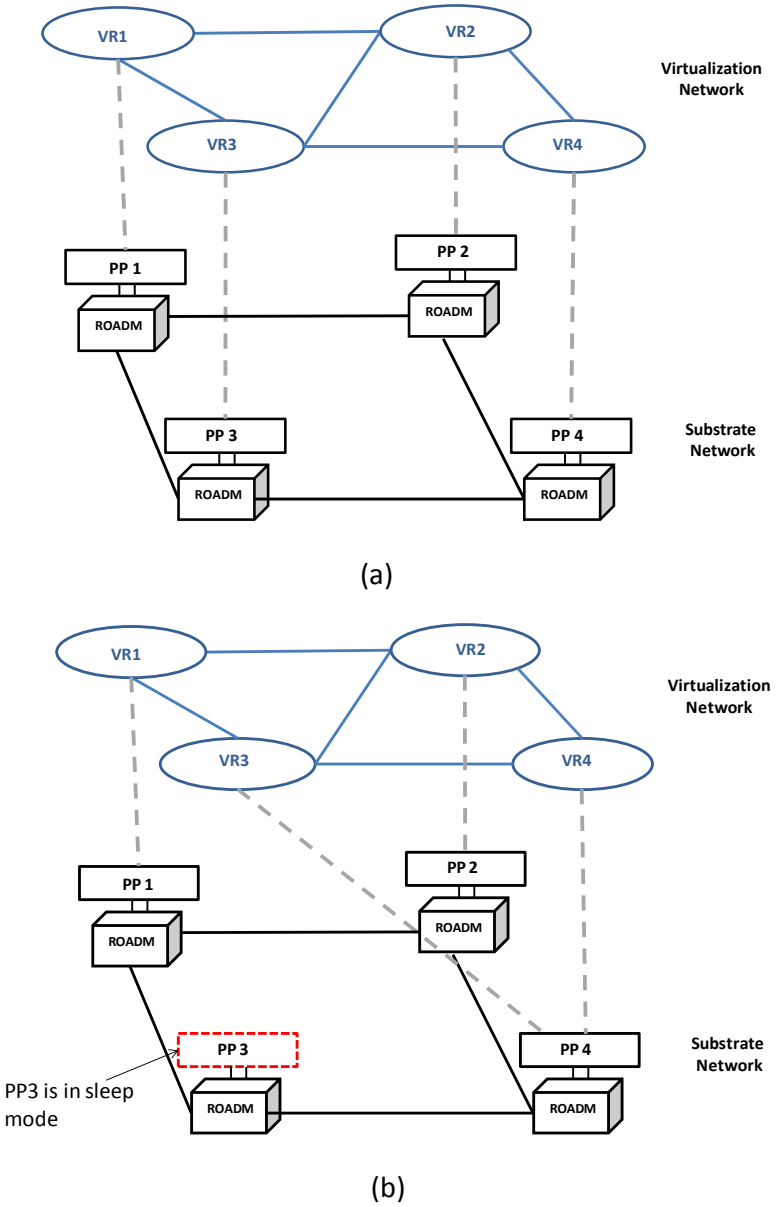

Figure 15. (A) Normal Network Configuration. (B) Energy Efficient Configuration where PP3 is in a Sleep State and VR3 and VR4 are running on PP4.

3.4 **Virtual Router Migration**

Our dynamic energy management framework leverages a recent advance in live virtual router migration called Virtual ROuter On the Move (VROOM) [44], which

allows a VR instance to freely move between PPs avoiding unnecessary changes to the logical topology. VROOM is transparent to routing protocols and results in no performance impact on the data traffic. Thanks to VROOM, our framework can improve the network energy efficiency without introducing traffic disruption or protocol reconvergence events.

In our work, we are not interested in accelerating or improving the VROOM scheme. Instead, VROOM is treated as a subroutine when moving multiple VRs to their appropriate PP locations in a single VRM. In this section, a single VRM is described briefly and subsequently a multiple VRM procedure is presented. The single VRM procedure is based on the VROOM system in an optical core network environment. The purpose of exploring a single VRM scenario is to examine the procedures used in the multiple VRM scenario. Finally, the migration time is discussed.

3.4.1 **Single VRM Procedure**

The live virtual router migration technology VROOM [44] has been introduced in Section 2.3.3. Based on VROOM, a single VRM procedure for our network architecture is shown as follows:

1. Set up a temporary tunnel between PP_{cur} and PP_{dest}. The tunnel is used to transmit the VR's Control Plane (CP) state information to the new location and re-direct routing messages during the link migration process (as described in Step 4).
2. The CP state is copied and transmitted from PP_{cur} to PP_{dest} via the established temporary tunnel. When CP state transmission finished, the migrated CP functionality starts exchanging routing messages via the tunnel established in Step 1to the original DP entity with its neighbours and the old / original CP functionality ceases operation.
3. Then, instead of transmitting the original DP state from PP_{cur} to PP_{dest}, the DP functionality is re-instantiated by the migrated CP entity on PP_{dest}. It takes a short period of time (e.g. 20 seconds) to install the new DP functions on PP_{dest}. During

this time, traffic is still forwarded by the old DP entity on PP_{cur}. The CP entity on PP_{dest} still acts as a remote CP entity.

4. After the new DP state information repopulation is complete, links can be moved from PP_{cur} to PP_{dest}. This procedure is called link migration. In the optical layer, this can be achieved efficiently by dynamically reserving and releasing optical channels. When new optical links are established from a migrated VR and its neighbours, the data traffic starts being transported along the new links. During this period, some traffic is still transmitted on the old links. Hence, the old links are not released immediately. Instead, a timer is set for these old links, which is slightly larger than the time needs to transport the traffic through the longest path in the network. When the timer expires, the old links are released and the old PP may be put into the sleep state (if there is no VR running on it). The penalty of this mechanism is that the energy is not saved during the hold-down period of timer.
5. Finally, data traffic via the old links are complete, the original DP entity on the old location and the temporary tunnel as well as the old links themselves can be removed.

In summary, a single VRM can be divided into two stages: CP migration and link migration. The CP migration stage includes migrating CP operations and new DP repopulation. On the other hand, the link migration stage consists of establishing new links for the migrated VR and data traffic flows through the new links, waiting for the traffic on old links to complete and removing the old links, the tunnels and the old DP state/functions.

3.4.2 **Multiple VRM Procedure**

In our framework, one or multiple VR(s) need to be migrated if VRM conditions are satisfied. The number of migrated VRs is decided by the output of the destination physical platform selection algorithm (as discussed in Section 3.5 and Chapter 4) and the underlying optical resource availability (as discussed in Section 3.7).

Based on the single VRM procedure described in Section 3.4.1, there are two main methods of moving multiple VRs in our framework. Firstly, this can be achieved by repeating the single VRM mechanism several times according to the number of VR instances to be migrated. This means that a VR starts moving when the previous VRM has finished. However, this method has two drawbacks. Firstly, it is wasteful to set up and remove links several times during the consecutive VRM procedures. For example, a migrated VR needs to re-establish links to its neighbours from the new location. Assuming one of the migrated VR neighbours also needs to be moved to a new location, then some newly established links in previous migration may need to be removed and some new links need to be set up between the migrated VR and its migrated neighbour's new location. The *establish-remove* link process may repeat many times until multiple VRM completes, which is wasteful and unnecessary. Secondly, moving several VRs one by one takes a long time, which is not efficient. Hence, it is not desirable to move VR instances in a succession of VRM events.

The second method to realize multiple VRM is to movie VR instances concurrently, which is more efficient. According to the two stages, i.e.CP state migration and link migration, in the single VRM procedure, it is not difficult to execute the CP migration for all migrated VR instances at the same time. This means that some temporary tunnels are established between source PPs and destination PPs. Then, all CPs of migrated VRs can be moved to their new locations to repopulate DP entities. However, the challenging part is how to set up new links for migrated VRs and their neighbours if some neighbours also move to new locations during the link migration stage. In this case, migrated VRs need to know their neighbours' updated location information (no matter whether they remain at their old locations or move to new locations). Fortunately, since a CCU is used in the framework and contains all the migration information, it can inform the migrated VRs of their neighbours' updated location information. For instance, when the CP functionality of a migrated VR is copied to a new location, it can establish new links to its neighbours based on the CCU messages. When the migrated VRs finish

establishing new links with their neighbours, data traffic can start flowing through the new links and the old links can be removed.

The latter approach requires sharing more information with VRs during the link migration process compared to the VROOM mechanism. Thus, whilst it is generally advantageous to move multiple VR instances concurrently, the increasing system complexity for coordinating link migration may offset the benefits.

3.5 **Destination Physical Platform Selection**

An important question with a dynamic energy management framework is where to move VR instances to. We call it a Destination Physical Platform Selection (DPPS) problem. As network traffic demand has a well-known regular daily pattern caused by human activities [40] [41], the physical network composed of active physical platforms (PPs) can be expanded and contracted according to the varying traffic demand. Some VRs can be moved to a smaller set of PPs when the traffic demand is light. Much of the functionality in the unused PPs can then be put to sleep to save energy. As traffic demand increases some sleeping PPs can be re-awoken and the VR functionality returned to these PPs close to where the traffic load is rising. There are multiple constraints that need to be considered in a DPPS problem as follows:

1. A source PP and a destination PP must be compatible with each other. If two PPs are not compatible, a VR may not be able to run on the destination PP. PPs (physical routers) produced by various vendors generally employ their own operating systems. If PPs come from different vendors, a "translator" or shim layer would be needed to enable a VR to work on dissimilar PPs, which is beyond the scope of this thesis. We assume that all PPs use same operating system and are compatible, though they need not be identical.
2. A destination PP must have sufficient capacity to accommodate a new VR instance without detrimentally impacting on the performance of any VR instances it is

currently hosting. For example, a destination PP must have enough bandwidth to handle the traffic load associated with the new VR instance as well as the sufficient processing power.

Fortunately, these two constraints are not difficult to be satisfied based on some current ISP network features. Firstly, most ISPs generally use physical routers from one or two vendors in order to simplify the network maintenance and decrease operational costs. Hence, this provides plenty of opportunities that a large number of physical routers can be selected as a destination PP. Then, ISPs typically design their network with over-provisioning in order to sustain the maximum traffic demand or some unexpected events. This feature may enable a PP to accommodate several VR instances.

In summary, DPPS is a complex challenge which can be formulated as an optimization problem whose objectives are to maximize the energy saving and minimize the migration cost whilst considering two constraints above. This problem is discussed in detail in Chapter 4.

3.6 **Triggering Virtual Router Migration**

Another important question to be answered is when to trigger the VRM. The question is also about whether the traffic can be predicted. If network traffic can be predicted, a proactive mechanism can be used which allows a slow and complex algorithm to obtain the optimal result for maximizing the energy saving and minimizing the migration cost. Otherwise, a reactive and fast mechanism needs to be used.

In order to explore the network traffic characteristics, a traffic analysis study has been undertaken as described in Appendix A using historical traffic data from the Abilene network. According to the data provided, results show that the traffic between some source-destination pairs can be predicted relatively well whilst this may not be the case in other instances. The traffic prediction problem is more complex than might be

initially assumed. Therefore, given traffic forecasting limitations a reactive mechanism was used in our framework.

3.7 **Dynamic Optical Connection Management**

A key constraint of the framework is to maintain the logical IP-layer topology unchanged whilst consolidating VRs onto smaller set of PPs to save energy when appropriate. According to the discussion in Section 3.4, it is possible to use VROOM to avoid unnecessary changes to the logical topology as well as avoid a performance impact on the data traffic. However, the VROOM system does not describe events in the optical layer. In this section, a dynamic optical connection management scheme is discussed in detail.

In the optical layer, the routing of optical connections is typically based on a shortest path routing algorithm in order to minimize the usage of optical resources. Three types of optical connections are established and released dynamically in our framework.

1. New optical connections are established between $PP_{current}$ and PP_{dest} to transmit the CP state of VR instances. These optical connections are then released when VRM completes.
2. New optical connections are established between $PP_{default}$ and PP_{dest} if a VR is moved away from its default PP location. These optical connections are used for forwarding the traffic to the remotely located VR instance that performs the packet processing and exchange of signalling messages.
3. New optical connections are established between the migrated VRs with their neighbours. The original optical connections between the migrated VRs' original location and their neighbours are released when VRM completes.

Figure 16 provides a simple example illustrating how and why the architecture places additional demands on the optical connectivity resources. Figure 16(a) shows the path taken by traffic from host A to host B before VRM which passes through VR 1, 2,

3, 4 on PP A, B, C, D, respectively. Assuming the traffic volume of PP A is relatively low, VR 1 is moved to PP C and then PP A is allowed to sleep. Consequently, VR 1 and VR 3 are hosted in the same PP. To ensure VR 1 and VR 3 remain logically separated from each other, extra optical connections are added. Conversely, if this was not the case, and VR 1 and VR 3 were made visibly adjacent to each other at the IP layer, the logical topology would be affected. By permitting information to flow directly between them, they would exchange routing information and thus result in a change leading to a reconvergence event.

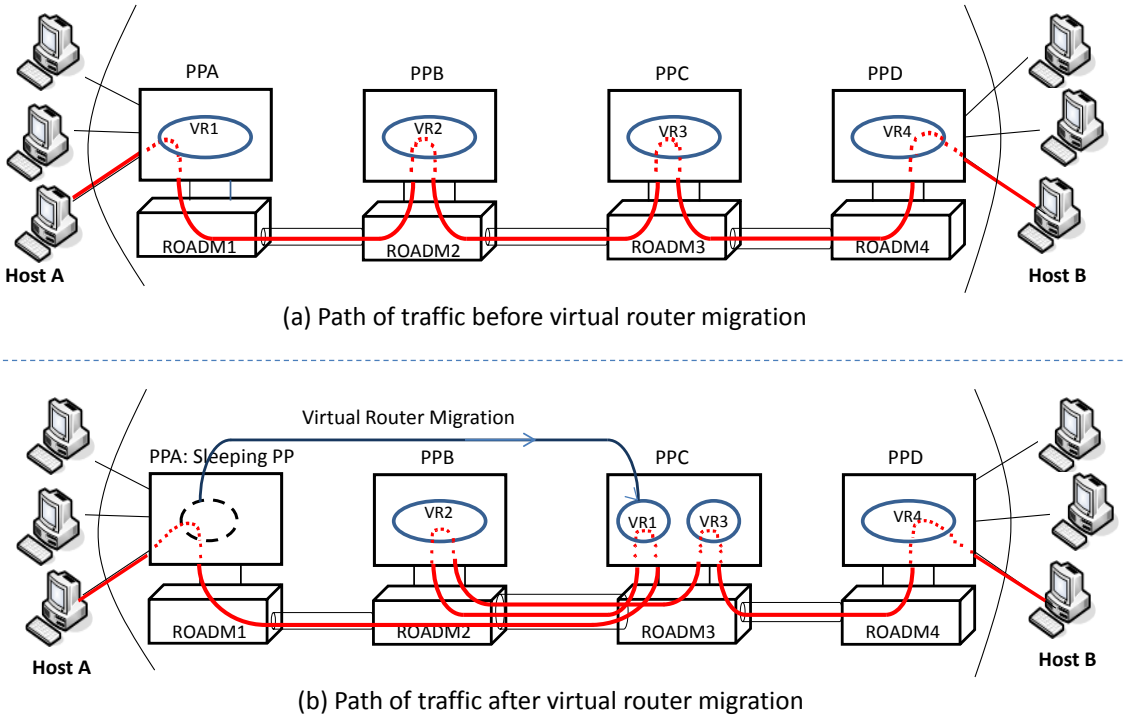

Figure 16. Example Scenario (a) Before and (b) After Virtual Router Migration showing the Need for Additional Optical Connections

Figure 16(b) shows that one optical connection now forwards traffic from the customer (Host A) directly to the new location of VR 3 without passing through the Layer-3 functionality of PP B. A new connection is then used to send traffic from VR 1 to VR 2. The location and number of optical connections that must be created is

determined by the location of where the VR migrates to and the number of Layer-3 interconnections that the VR possesses. Therefore, a migration to a remote location imposes an associated cost by requiring additional optical connections when the VR is moved to the destination PP. This is a multi-commodity problem as although it is beneficial to pack multiple VR instances onto the same PP, the additional optical resources it consumes offsets this benefit.

3.8 Network Power Consumption Model

In this section, a component-based power consumption model is introduced with reference to the network architecture described in Section 2.2.1. The power consumption models of important devices: Physical Platforms (PPs), Reconfigurable Optical Add Drop Multiplexers (ROADMs) and Optical Line Amplifiers (OLAs) are firstly introduced in Section 3.8.1 and then the total network power consumption model is summarized in Section 3.8.2.Finally, a cumulative energy consumption model is described in Section 3.8.3.

In theory, optical lasers could be switched on/off based on the channel requirements. However, every time when one/more lasers are switched on / off, power equalization is required which consumes energy and takes time. Hence, in our study, we do not consider switching optical lasers and assume that ROADMs as well as OLAs remain working throughout the whole day.

3.8.1 Individual Equipment Power Consumption Model

It is important to note that some functions of PPs and ROADMs have not been fully realized in existing commercial products (as explained in Section 2.2.2 and 2.2.3), e.g. sleep function. Thus, publicly available product data-sheets are not able to provide suitable energy consumption values because these functions are not considered. When we build a power consumption model, we consider both the ideal architecture,which

contains equipment possessing a complete set of the required functionalities, and an existing commercial product architecture which provides some estimated values of power consumption. Note that the unit of power consumption of all components and devices is Watts.

1. Physical Platform

The power consumption model of a PP is based on publicly available data sheets of one of major commercial core routers: the Cisco Carrier Routing System (CRS) series [163]. Besides the conventional single-shelf router architecture, the Cisco CRS series also support a multi-shelf routing system which employs modular and distributed architectures to provide very high capacity, e.g. up to 92Tbps in CRS-1 series and up to 322Tbps in CRS-3 series [164]. The power consumption models of single-shelf and multi-shelf routing systems are introduced as follows.

Firstly, a single-shelf PP is basically made up of a base system and a line-card block. The *base system* includes a router processor, a switch fabric and some functional modules, e.g. power, alarm and cooling. In a single-shelf CRS, the base system is called a Line Card Chassis (LCC). Specifically, we do not consider power consumption to depend upon traffic utilization in the base system. This is because the power variation related to traffic utilization is only about 10% of the total in existing commercial products, which is less significant than the power consumption of the line-cards [25]. Meanwhile, the *line-card block* includes multiple line-card slots. In the CRS system, a single LCC can support up to 16 line-cards. Thus, the power consumption of a single-shelf physical router / PP is simplified to being a sum of a base system, P_{base}, and a number of installed line-cards [22].

$$P_{single} = P_{base} + n_{lc} \cdot P_{lc} \tag{3.1}$$

where n_{lc} is the number of installed line-cards on the PP. P_{lc} represents the power consumption of a line-card. The base system is called a LCC in a Cisco single-shelf router. Thus, Equation 3.1 can be re-written as:

$$P_{single} = P_{lcc} + n_{lc} \cdot P_{lc} \quad (3.1a)$$

where $P_{lcc}$is the power consumption of a LCC.

Generally, a single-shelf routing system cannot support high-capacity core networks because the number of line-card slots is limited. Thus, router manufacturers have developed multi-shelf routing systems that employ modular and distributed architectures in order to provide high capacity in a single system. For example, a Cisco multi-shelf routing system uses a component called a Fabric Card Chassis (FCC) to interconnect multiple LCCs. Meanwhile, a FCC can connect up to 9 LCCs. Note that a LCC and a FCC both belong to the base system.

In the power consumption model of a multi-shelf routing system, both the number of LCCs and FCCs needs to be considered. The number of LCCs, n_{lcc}, is determined by the total number of installed line-cards, n_{lc}, and the maximum number of line-cards each LCC supports, M_{lc}. Thus, the number of LCCs is represents as:

$$n_{lcc} = \left\lceil \frac{n_{lc}}{M_{lc}} \right\rceil \quad (3.2)$$

Similarly, the number of FCCs is determined by the total number of installed LCCs, n_{lcc}, and the maximum number of LCCs that each FCC can support, M_{lcc}, as follows:

$$n_{fcc} = \left\lceil \frac{n_{lcc}}{M_{lcc}} \right\rceil \quad (3.3)$$

Thus, the power consumption of a multi-shelf PP is represented as:

$$P_{multi} = P_{base} + n_{lc} \cdot P_{lc}$$

$$= n_{lcc} \cdot P_{lcc} + n_{fcc} \cdot P_{fcc} + n_{lc} \cdot P_{lc} \qquad n_{lc} > M_{lc} \quad (3.4)$$

Where $P_{fcc}$is the power consumption of a FCC. Then, we obtain $n_{lcc}$and n_{fcc} from Equation 3.2 and 3.3, so Equation 3.4 can be re-written as follows:

$$P_{multi} = P_{base} + n_{lc} \cdot P_{lc} \left\lceil \frac{n_{lc}}{M_{lc}} \right\rceil \cdot P_{lcc} + \left\lceil \frac{\left\lceil \frac{n_{lc}}{M_{lc}} \right\rceil}{M_{lcc}} \right\rceil \cdot P_{fcc} + n_{lc} \cdot P_{lc} n_{lc} > M_{lc} \quad (3.4a)$$

Based on Equation 3.1 and 3.4, the power consumption of an active PP, P_{ppw}, is as follows:

$$\begin{cases} P_{ppw} = P_{single} = P_{lcc} + n_{lc} \cdot P_{lc} & n_{lc} < M_{lc} \\ P_{ppw} = P_{multi} = \left\lceil \frac{n_{lc}}{M_{lc}} \right\rceil \cdot P_{lcc} + \left\lceil \frac{\left\lceil \frac{n_{lc}}{M_{lc}} \right\rceil}{M_{lcc}} \right\rceil \cdot P_{fcs} + n_{lc} \cdot P_{lc} & n_{lc} > M_{lc} \end{cases} \quad (3.5)$$

Moreover, we assume that an entire PP can be put into sleep or active states dynamically depending on the network requirements. For instance, the entire PP can be put into a sleep state provided it is not hosting any VR instances. Thus, when a PP is in the sleep state, the base system and line-cards stop working except for part of the management module within the base system. The management module is maintained for exchanging the signalling messages, which consumes a small fraction θ of the active base system power consumption. As a result, the power consumption of a sleeping PP P_{pps} is:

$$P_{pps} = \theta \cdot P_{base} \quad (3.6)$$

According to Equation 3.5 and 3.6, the total power consumption of PPs, P_{pp_t}, in an N-node network is:

$$P_{PP\_t} = \alpha \cdot P_{ppw} + (N - \alpha) \cdot P_{pps} \quad (3.7)$$

Where α is the number of active PPs.

2. Reconfigurable Optical Add Drop Multiplexer

An ideal ROADM can provide a flexible way to add, drop or switch any wavelength to any direction. However, the current ROADM generation is not fully flexible (as explained in Section 2.2.3). In practice, the power consumption of an ROADM is determined by a node degree and an add / drop degree [109]. Meanwhile, the add / drop degree can be smaller than the node degree if some directions do not require the add / drop function. For example, the power consumption of an ROADM with 40 channels, P_{roadm}^{40}, is represented as follows [109]:

$$P_{ROADM}^{40} = \omega_1 \cdot 85 + \omega_2 \cdot 50 + 150 \qquad \omega_1 \geq \omega_2$$

where ω_1 is the node degree and ω_2 is the add / drop degree.

As we require any wavelength to be added or dropped at any direction, the node degree needs to equal the add / drop degree. The equation can therefore be re-written as follows:

$$P_{ROADM}^{40}(\omega) = \omega \cdot 135 + 150$$

where ω is the node degree. Thus, the power consumption of an ROADM can be expressed as $P_{ROADM}(\omega)$. In addition, we assume that all ROADMs remain active in order to support traffic transmission in the optical layer.

The total power consumption of ROADMs, P_{ROADM_t}, in an N-node network is:

$$P_{ROADM\_t} = \sum_i^N P_{ROADM}(\omega_i) \tag{3.8}$$

3. Optical Line Amplifier

An OLA is a device that is used for amplifying an optical signal directly without the need for O-E-O conversion. An OLA is typically unidirectional and a bidirectional OLA is composed by two unidirectional OLAs in practice. The total power consumption of OLAs, P_{OLA_t}, in an N-nodes network is:

$$P_{OLA\_t} = \sum_{i,j;i\neq j}^N \left\lceil \frac{d_{i,j}}{\varepsilon_{OLA}} \right\rceil \cdot e_{OLA} \tag{3.9}$$

where $d_{i,j}$ represents the physical link length between node i and node j and e_{OA} is the powerconsumption of a bidirectional OLA. ε_{OLA} is the maximum allowed link length without the need of amplification. A typical value of ε_{OLA} is 80km. Additionally, we assume that the power consumption of an OLA does not depend on the traffic and each OA amplifies the entire C-band.

4. Centralized Control Unit

We assume that the CCU can be a computer that operates the functionality as described in Section 3.2.1. In the simulation, a desktop computer is used and the

configuration of the computer is described in Section 6.2. As the simulation results show in Section 5.6, the computer is able to compute the destination physical platform selection algorithm in tens of seconds which is satisfactory, given the monitoring interval. The desktop computer typically consumes around 50 Watts [192], which is trivial compared with a physical router or ROADM. Hence, we neglect the energy consumption of the CCU.

3.8.2 **Total Power Consumption Model**

In the framework, every node in the substrate network is composed of a PP and an ROADM. Based on the Equation 3.7 -3.9, the total power consumption P_{total} of an N-nodes network is thus:

$$P_{total} = P_{PP\_t} + P_{ROADM\_t} + P_{OLA\_t} \tag{3.10}$$

3.8.3 **Cumulative Energy Consumption Model and Energy Efficiency**

In this section, two cumulative energy models are described, i.e. a Base-line scenario without VRM and an Energy-efficient scenario with VRM. Based on these two models, the energy efficiency of the framework can be computed.

Firstly, we assume that the overall duration is T and there are N nodes in the network labelled from 1 to N. Meanwhile, the initial power consumption of a node i is P_i^0. In the base-line scenario without VRM, the cumulative energy consumption is as follows:

$$E_{base\_line} = \sum_{i=1}^{N} P_i^0 \cdot T \tag{3.11}$$

Then, in an energy efficient scenario, we assume that there are K VRM events and the time of jth VRM event is denoted as t_M^j. At the same time, after the $j$thVRM, the power consumption of the node i may change and it is denoted by P_i^j. The relationship

between the migration time and the node power consumption after the migration is shown in Figure 17.

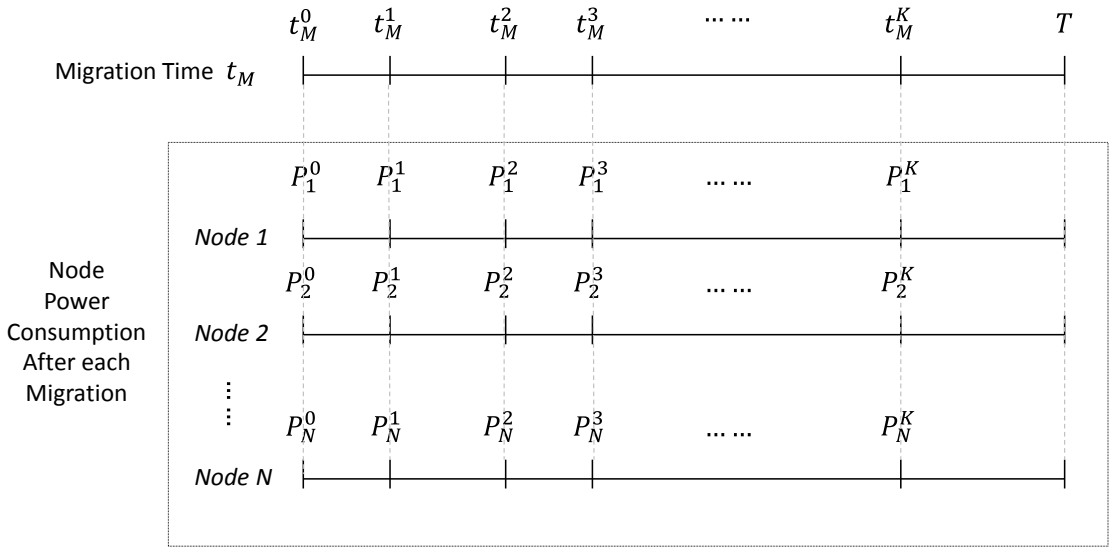

Figure 17. Relationship between Migration Event Time and the Node Power Consumption after Migration

The cumulative energy consumption for a node i can be represented as:

$$E_{VRM}^{i} = P_i^0 \cdot (t_M^1 - t_M^0) + P_i^1 \cdot (t_M^2 - t_M^1) + \cdots + P_i^{K-1} \cdot (t_M^K - t_M^{K-1}) + P_i^K \cdot (T - t_M^K)$$

$$= \sum_{j=1}^{K} P_i^{j-1} \cdot \left(t_M^j - t_M^{j-1}\right) + P_i^K \cdot (T - t_M^K) \quad (3.12)$$

Thus, the total cumulative energy consumption, E_{VRM}, in an N-nodes network is:

$$E_{VRM} = \sum_{i=1}^{N} \sum_{j=1}^{K} P_i^{j-1} \cdot \left(t_M^j - t_M^{j-1}\right) + P_i^K \cdot (T - t_M^K) \quad (3.13)$$

According to Equation 3.11 and 3.13, the percentage energy saving of the framework is presented as follows:

$$\text{Energy } saving\ \% = \frac{E_{base\_line} - \ E_{VRM}}{E_{base\_line}} \times 100 \quad (3.14)$$

3.9 **Summary**

In this chapter, a novel dynamic energy management framework is proposed. A key constraint of the framework is to maintain the logical IP-layer topology unchanged whilst consolidating VRs onto a smaller set of PPs to save energy, when possible. Hence, Virtual Router Migration (VRM) and the turning off of some unneeded Physical Platforms (PPs) are combined with an automatic optical layer management scheme to enable resources to be used in an efficient manner. In order to realise this new energy management framework, some requirements of the network architecture and associated functionalities are firstly described such as a centralized control unit for collecting the network state information and a PP and its components to possess a sleep function. Then, the overall dynamic energy management procedure is illustrated followed by an examination of several important issues such as the VRM procedure, where to move VR instances to and when to move VR instances.

CHAPTER 4 DESTINATION PHYSICAL PLATFORM SELECTION ALGORITHM

4.1 Introduction

In Chapter 3, the crucial issue of where to move virtual router instances to in a dynamic energy management framework (i.e. destination physical platform selection) is introduced. In order to solve this problem, we develop an evolutionary-based algorithm called Virtual Router Migration Multi-Objective Evolutionary Algorithm (VRM_MOEA) in Chapter 4. The remainder of Chapter 4 is organized as follows. Initially, Section 4.2 describes the destination physical platform selection problem. Then, some reasons why we use an evolutionary-based algorithm are given in Section 4.3. Next, the proposed algorithm VRM_MOEA is detailed in Section 4.4. Finally, a chapter summary is given in Section 4.5.

4.2 Problem Description

An important consideration in the dynamic energy management framework is how to select appropriate Physical Platforms (PPs) to move Virtual Router (VR) instances to, given various constraints. We call a Destination Physical Platform Selection (DPPS). DPPS is used to determine which VR instances are viable candidates to be migrated so that some PPs can be placed in their sleep state during off-peak hours, or to determine which PPs need to be re-awoken ready to accommodate VR instances during busy periods.

DPPS can be represented as follows. Let $\boldsymbol{P} = (\boldsymbol{PP}, \boldsymbol{E})$ be a substrate network where $\boldsymbol{PP}$ is a set of PPs, $\{PP_1, PP_2, \dots, PP_N\}$, and $\boldsymbol{E}$ is a set of physical links. Let $\boldsymbol{V} = (\boldsymbol{VR}, \boldsymbol{L})$ be a virtualized network where $\boldsymbol{VR}$ is a set of VR instances, $\{VR_1, VR_2, \dots, VR_N\}$, and $\boldsymbol{L}$ is a set of virtual links. Let $\boldsymbol{map}: \boldsymbol{VR} \rightarrow \boldsymbol{PP}$ be a function that maps the set of VRs onto the set of PPs. For instance, $\boldsymbol{map}(VR_1) = PP_2$ represents that VR_1 is running on PP_2.

An initial mapping $\boldsymbol{map_0}: \boldsymbol{VR} \rightarrow \boldsymbol{PP}$ represents each VR instance is running on its default PP. For a Virtual Router Migration (VRM) event, the current mapping (before the migration) is $\boldsymbol{map_c}: \boldsymbol{VR} \rightarrow \boldsymbol{PP}$ and the destination mapping (after the migration) is $\boldsymbol{map_d}: \boldsymbol{VR} \rightarrow \boldsymbol{PP}$. The aim of DPPS is to determine an efficient method to assign VR instances onto appropriate PPs. The mapping of the virtualized network and the substrate network is changed from $\boldsymbol{map_c}$ to $\boldsymbol{map_d}$. When the network condition satisfies any migration conditions, VR instances are assigned to PPs appropriately, satisfying objective functions and various constraints. Migration conditions, objective functions and constraints of DPPS are described as follows.

- *Migration Conditions*

In order to determine the appropriate migration conditions, relevant terms are defined in Section3.3. A brief review is provided as follows. A PP has two states: *Active* and *Sleep*. An active PP has three operating modes: *Quiet*, *Normal* and *Busy*. Two thresholds: *Quiet* and *Busy* are defined to distinguish between the three modes. *Quiet* and *Busy* thresholds are used as a condition to trigger a VRM event; they also provide hysteresis to prevent unnecessary frequent transitions. Based on these concepts, a VRM event is triggered in the following conditions:

1. If some PPs are in the *Quiet* mode, and are effectively underutilized, it is preferable to consolidate their VR instances on fewer remaining PPs. The unneeded PPs can be then put into their *Sleep* state to save energy.

2. Some sleeping PPs are selectively re-awoken when working PP(s) cross their *Busy* threshold. Then, VR instances are selectively moved away from PPs where traffic load is increasing to an unsustainable level.

- *Objective Functions*

In DPPS, the two objective functions that need to be optimized are described below.

1. The overall network power consumption is required to be minimized. Maximizing the consolidation of VR instances onto as few PPs as possible for a given traffic demand without degrading the network performance is preferred. Consequently, surplus PPs can be placed in the *Sleep* state to save energy.
2. The cost associated with VRM is required to be minimized. We also define three terms: a default PP ($PP_{default}$), a current PP (PP_{cur}) and a destination PP (PP_{dest}) in Section3.3. They are used to represent different PP locations of VR instances in the network. In DPPS, we focus on two types of migration cost. The first type considers the additional optical resources from a $PP_{default}$to its PP_{dest}. The longer the distance between $PP_{default}$ and PP_{dest}, the larger is the associated cost. The second type of the cost takes into account the distance from PP_{cur} to PP_{dest}. If $PP_{cur}$and$PP_{dest}$are far away from each other, it requires using a longer optical channel to transmit the VR instance.

- *Constraints*

Some constraints need to be considered in DPPS as follows.

1. A $PP_{cur}$and a PP_{dest}must be compatible with each other. If two PPs are not compatible, a VR instance may not be able to run on the PP_{dest}. In our study, we assume that all PPs are compatible.
2. A PP_{dest} should be able to accommodate new VR instance(s) without detrimentally impacting the performance of any VR instances it is currently hosting.

3. The optical resources must be able to support the new network configuration. For the wavelength assignment task, although all-optic wavelength conversion has been discussed for 20 years, it has not yet led to practical products that could be deployed in commercial networks [193]. Hence, the wavelength continuity constraint is considered in our scheme so O-E-O conversion is not permitted for pass-through optical connections. This requires sufficient optical channels of the same wavelength to transit traffic flows to remote VR(s); if not, the migration cannot take place. In addition, if all-optic wavelength conversion is possible, it is beneficial for the scheme as it removes a constraint on the optical layer.

Finally, based on the concepts of a MOOP in Section 2.3.5, a decision vector (a solution) is $\boldsymbol{l} = \{l_1, l_2, \dots, l_N\}$, where l_i represents a PP location of the VR_i in DPPS. An objective vector is $\boldsymbol{o} = \{f_1(\boldsymbol{l}), f_2(\boldsymbol{l})\}$, where $f_1 = \min(power\ consumption)$ and $f_2 = \min(VRM\ cost)$, given various constraints.

4.3 **Why MOEA**

According to the problem description in Section 4.2, an algorithm is required to meet the following conditions:

1. The algorithm is able to obtain reasonable solutions across different scenarios, i.e. during off-peak hours and peak hours. Generally, the two objectives of a potential solution are conflicting. For example, a solution that saves large amount of energy may also have a large VRM cost. This is due to that the consolidation of VR instances onto fewer PPs may explicitly increase VRM cost.
2. The algorithm is able to provide a reasonable solution in a relatively short time, such as 5 minutes or less. If the computation is slow, the network conditions may change significantly whilst a solution is being sought. In addition to the algorithm computation time, another period of time is also needed to complete the VRM procedure. Hence, a quick algorithm is preferred.

3. The algorithm can generate a group of solutions that provides more opportunities for VRM. The optical resource availability is not considered in the DPPS algorithm due to the computational complexity. However, the optical resource availability limits the possibility of VRM. Hence, having a group of solutions may provide more chances for a viable VRM.
4. The algorithm is able to search a large potential solution space efficiently. Although the number of candidate solutions is finite if the size of network is known, the search space can be large. The number of candidate solutions is N^N without considering constraints in an N-node network. For instance, a solution space of a 10-node network can consist of 10^{10} individuals. It is time-consuming work to examine all candidate solutions. Hence, it is preferential to find algorithm that can search efficiently.

Considering the four requirements above, an algorithm based on Multi-Objective Evolutionary Algorithms (MOEAs) called Virtual Router Migration – MOEA (VRM_MOEA) is developed. The definitions and principles of Evolutionary Algorithms (EAs) and MOEAs are addressed in Section 2.3.5. A brief introduction is also provided as follows.

EAs typically start with a set (population) of candidate solutions generated by a random method or a predefined scheme. A candidate solution is also called an individual or a chromosome. In this chapter, we use the term candidate solution, individual and chromosome interchangeably. The population is then refined iteratively by employing two basic principles: selection and reproduction. The selection mechanism mimics the fierce competition for survival in the natural world. The fitter individuals survive and have a higher chance of producing offspring. The poor quality individuals are usually eliminated from the population. A reproduction mechanism, including mutation and crossover genetic operators, imitates the process of producing offspring, in the hope of producing individuals with better fitness (quality). The selection and reproduction operations repeat until a stopping criterion is reached, such

as the maximum number of iterations (generations) has been reached or the quality of solutions stops increasing.

MOEAs are typically used for solving problems involving multiple conflicting objectives by EA mechanisms. As several objectives are conflicting, there is no single solution which is able to simultaneously optimize all objectives. Hence, the outcome of multi objective optimization problem (MOOP) is a set of "trade-offs" or compromises. As we described in Section 2.3.5.1, there are many methods to solve MOOPs. Compared with other methods, MOEAs are able to find a diverse set of solutions for difficult problems with non-convex, discontinuous solution-spaces as MOEAs are population-based methods. MOEAs can probably obtain (or approximate) Pareto-optimal solutions in a single optimization run instead of obtaining one solution each run [147]. In addition, MOEAs usually do not require weight, scale or prioritized objectives [147]. Therefore, MOEAs are one of the most powerful mechanisms to solve the multi-objective problems (often with conflicting goals) and have been successful applied to a wide range of practical problems [148]. Note that MOEAs provide a generic methodology and they need to be customized to the particular problem.

Based on the advantages of MOEAs, we develop an algorithm called VRM_MOEA using a MOEA methodology. The VRM_MOEA algorithm is described in detail in Section 4.4.

4.4 **Destination Physical Platform Selection Algorithm-VRM_MOEA**

In this section, the details of the destination physical platform selection algorithm – VRM_MOEA are presented. Initially, the parameter design is described in Section 4.4.1, which provides the notations for variables and counters in the algorithm. Next, two important issues, i.e. a chromosome encoding and a viability test are introduced in

Section 4.4.2 and Section 4.4.3, respectively. Finally, the steps of VRM_MOEA are described in Section 4.4.4.

4.4.1 **Parameter Design**

The VRM_MOEA parameter design is described as follows.

1. N: The number of nodes in a network.
2. Let $\boldsymbol{P} = (\boldsymbol{PP}, \boldsymbol{E})$ be a substrate network where $\boldsymbol{PP}$ is a set of PPs, $\{PP_1, PP_2, \dots, PP_N\}$, and $\boldsymbol{E}$ is a set of physical links.
3. Let $\boldsymbol{V} = (\boldsymbol{VR}, \boldsymbol{L})$ be a virtualized network where $\boldsymbol{VR}$ is a set of VRs, $\{VR_1, VR_2, \dots, VR_N\}$, and $\boldsymbol{L}$ is a set of virtual links.
4. M_{lc}: the maximum number of line-cards of a PP.
5. χ: the percentage number of the number of line-cards, which is used for reserving line-card resources in a migration constraint.
6. R_{lc}: the required number of line-cards of a VR.
7. CAP_{PP}: the capacity of a PP.
8. TRA_{VR}: the traffic load of a VR.
9. PP_{cur} : a PP location where a VR is running on.
10. PP_{dest}: a PP location where is a (potential) destination of a VR after a VRM event.
11. $PP_{default}$: a PP location where is the initial location of a VR.
12. L_s: the length of a candidate solution/ a chromosome / an individual.
13. PP state: *Active* and *Sleep* state.
14. PP operating modes: *Quiet, Normal* and *Busy.*
15. Two thresholds of a PP: Busy threshold $T_{Busy}$and Quiet threshold T_{Quiet}. They are the percentage number of a PP capacity.
16. P_0: the initial primary population.
17. P_n: the primary population at generation n.
18. P_{size}: the size of a primary population.
19. S_0: an initial secondary population.

20. S_n: the secondary population at generation n .

21. S_{size}: the size of a secondary population.

22. M_n: the mating pool at generation n.

23. M_{size}: the size of a mating pool.

24. ρ_m: the mutation rate. $\rho_m \in [0,1]$.

25. ρ_c: the crossover rate. $\rho_c \in [0,1]$.

26. t: a generation counter.

27. T_{max}: the maximum number of generations.

4.4.2 **Chromosome Encoding**

In VRM_MOEA, a chromosome/an individual/a candidate solution is a string representation which indicates a possible mapping of VR instances onto PPs. There are several encoding choices in EAs, e.g. bit-string, floating points and integer representations. Because both VRs and PPs are labelled with a unique decimal number$\{1,2,\dots,N\}$in an N-node network, it is natural and straightforward to use an integer representation in VRM_MOEA.

In a chromosome, the particular position or *locus* is called a gene. The gene represents the index of a VR instance. For example, the first gene represents VR_1. The length of a chromosome is also determined by the number of VR instances. Moreover, the numeric value of a gene is referred to as *allele value* or *allele* which represents the index of a PP which hosts the corresponding VR instance.

An example of a chromosome is shown in Figure 18. The allele values of the chromosome are "121421". Firstly, the chromosome length indicates there are six VR instances and PPs in the network, assuming the default sate is for one VR to reside on one PP. However, in this case we can also see that these six VR instances are assigned onto three PPs, i.e. PP_1, $PP_2$and PP_4 because the allele values are 1, 2 and 4. VR_1, VR_3 and VR_6 are assigned ontoPP_1 because the 1st, 3rd and 6th genes have allele value 1. Similarly, VR_2 and VR_5 are assigned onto $PP_2$and VR_4 is assigned toPP_4. On the other

hand, as there is no 3, 5 and 6 appear in the chromosome, this means that PP_3, PP_5 and PP_6 can be put to sleep if the chromosome is selected as a new network configuration.

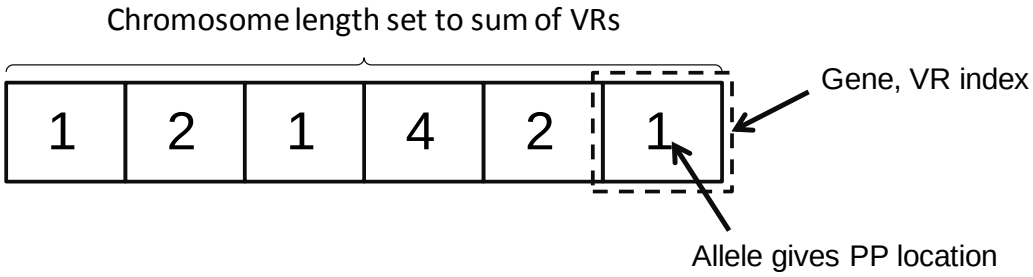

Figure 18. An Example Individual / Chromosome

Generally, a random method is used for generating chromosomes in the initial population so each allele value is randomly selected from the (i.e. 1 to N). However, this random method has a drawback that it does not consider migration constraints (as explained in Section 4.4.3). For example, a PP may be assigned several VR instances in the random method even if it does not have enough physical resource. Hence, in order to make sure that a chromosome satisfies various constraints, a viability test is applied to the chromosome. If a chromosome is not able to pass the test, it is eliminated from the population. The migration constraints and the viability test are explained in Section 4.4.3.

4.4.3 **Viability Test**

In VRM_MOEA, a viability test is required when a chromosome is created in order to satisfy two migration constraints, i.e. PP capacity and the number of PP line-cards. The details of these two constraints are described as follows.

- PP capacity

A destination PP should have enough resource to accommodate (potential) allocated VR instances. A PP cannot be considered as a destination PP if it will be in *Busy* mode

after accommodating several VR instances. Thus, a busy threshold, T_{Busy}, is used to guarantee this constraint.

Let n_{VR}^{j} be the number of VR instances which are assigned to a specific PP called PP_j and $\boldsymbol{VR_j} = \left\{VR_1, \dots, VR_{n_{VR}^{j}}\right\}$ be the set of these VR instances. Meanwhile, TRA_{VR}^{i} denotes the traffic load of a VR instance in $\boldsymbol{VR_j}$ and CAP_{PP}^{j} represents the capacity of PP_j. Hence, the PP constraint for PP_j is represented as:

$$\sum_{i=1}^{n_{VR}^{j}} TRA_{VR}^{i} < T_{Busy} \cdot CAP_{PP}^{j} \tag{4.1}$$

- Number of PP line-cards

The number of line-cards in a PP is another important constraint that needs to be considered in the viability test. The maximum line-card number of a specific PP called PP_j is denoted as M_{lc}^{j}. When several VR instances are assigned to the same PP, the sum of the required number of line-cards must not be larger than M_{lc}^{j}. However, in order to reserve some line-card resources to handle the traffic (if it increases), a safety margin, which is a percentage of the number of line-cards in a PP, χ, is introduced.

Similar with the PP capacity constraint, let n_{VR}^{j} be the number of VR instances which are allocated to PP_j. The maximum number of line-cards of PP_j is M_{lc}^{j}. Hence, this constraint can be represented as follows:

$$\sum_{i=1}^{n_{VR}^{j}} R_{lc}^{i} < \chi \cdot M_{lc}^{j} \tag{4.2}$$

where R_{lc}^{i} is the required number of line-cards of VR_i.Note that both Equation 4.1 and 4.2 are applied for all PPs which need remain working in a candidate solution.

The viability test is invoked when a new candidate solution is created, i.e. in the initial population generation and in the reproduction process for offspring at each

generation (as explained in Section 4.4.4 Step 1 and Step 6, respectively). A candidate solution can be put into the corresponding population only if it passes the viability test. Otherwise, it is eliminated. The *create- eliminate* process repeats until the corresponding population is full.

4.4.4 **VRM_MOEA Steps**

In this section, the steps of VRM_MOEA are described. Three populations: a primary population P, a secondary population S and a mating pool M are used in VRM_MOEA. The primary population is a regular population which is replaced by offspring every generation. The function of the secondary population is to archive nondominated solutions so no matter how many generations are iterated, the best nondominated solutions are retained. The mating pool is used for accommodating parents' chromosomes that produce offspring. The size of the three populations is constant in VRM_MOEA.

The algorithm flowchart is shown in Figure 19 and each step is described in detail from Step 1 to Step 7. The evolution of three populations at each generation is illustrated in Figure 20. The step number (i.e. from Step 2 to Step 6) along with the arrows indicates the sequence of the corresponding operation(s) amongst three populations. Step 1 works on the initial primary population and the secondary population and the output of VRM_MOEA is generated in Step 7.

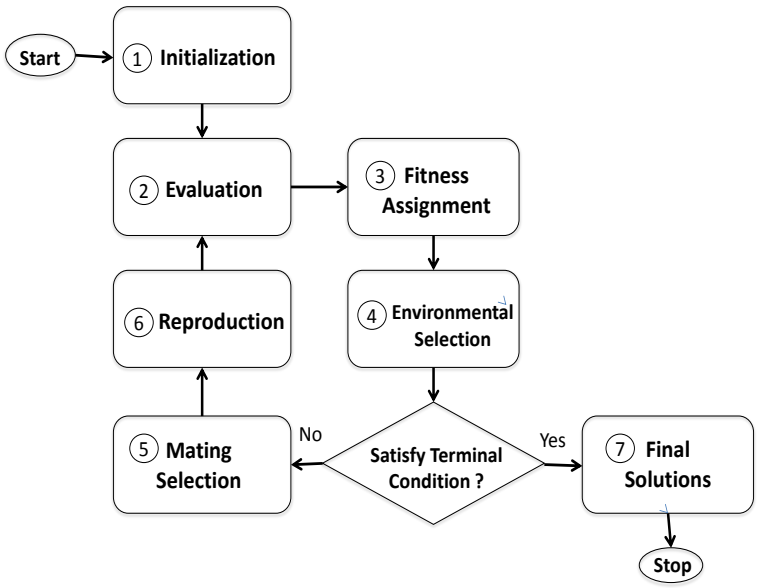

Figure 19. VRM_MOEA Flowchart

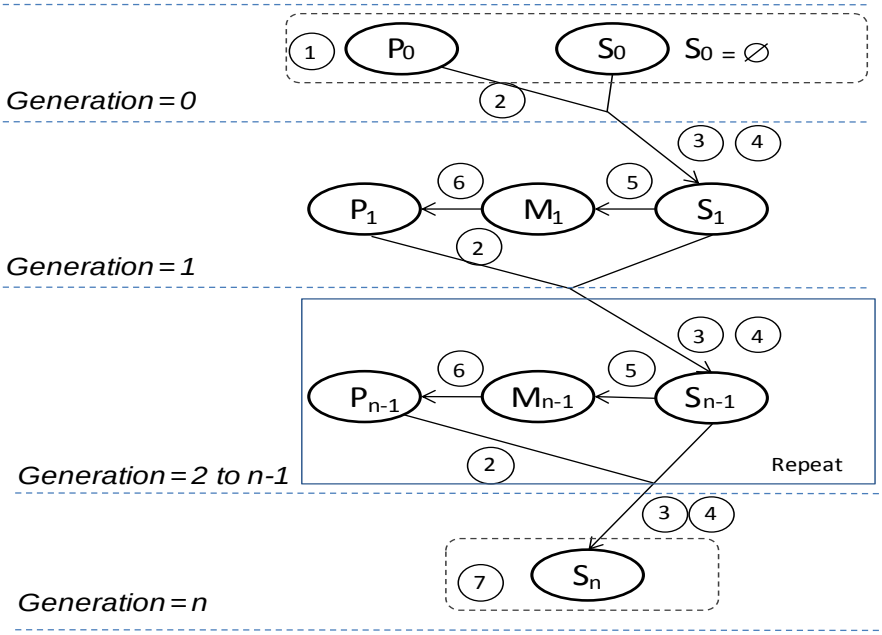

Figure 20. Population Evolution

Step 1: Initialization

At generation $t = 0$, an initial primary population P_0 and an empty secondary population S_0 are generated. P_0 has P_{size} candidate solutions. A new candidate solution

is generated randomly so it needs to be examined by the viability test (as explained in Section 4.4.3). If the new candidate solution passes the viability test, it is put intoP_0. Otherwise, it is eliminated and the generation process repeats until P_0 is full.

Step 2: Evaluation

In the evaluation stage, each candidate solution in the primary population is evaluated and then assigned two objective values. These two objective functions are described as follows.

- Power Consumption

The first objective function considers the power consumption of a candidate solution. The objective value is calculated by the power consumption model in Section 3.8. Because all ROADMs and OLAs remain working throughout the day, the function focuses on the PP power consumption.

- VRM Cost

The second objective function is the VRM cost. In this investigation, the VRM cost is considered in terms of the additional optical resources consumed that is associated with the cost during the VRM process and after VRM. For simplicity, VRM cost is measured in terms of hop-distance. In order to describe the hop-distance clearly, a function $dis(a, b)$ is used for representing the hop-distance between two PPs, i.e. PP a and PP b.

The VRM cost comprises two components. The first VRM cost component takes into account the additional optical resources from $PP_{default}$ to PP_{dest}. When a VR instance is not running on its $PP_{default}$, some additional optical connections are required to transmit traffic (as explained in Section3.7). Hence, the first migration cost component, $Cost\_a$, of an individual, i, is:

$$Cost\_a(i) = \sum_{j=1}^{N} dis\left(g^{j}_{default}, g_j\right) \tag{4.3}$$

where $g^{j}_{default}$ is $PP_{default}$ and g_{ji} is a potential PP_{dest} for VR_j.

The second component considers the cost within the VRM procedure. Because a temporary tunnel needs to be built to transmit the CP state of a VR instance, the second cost component calculates hop-distance from PP_{cur} to PP_{dest}. The second VRM cost component, $Cost\_b$, of an individual, i, is:

$$Cost\_b(i) = \sum_{j=1}^{N} dis\left(g^{j}_{current,}\, g_j\right) \tag{4.4}$$

where $g^{j}_{current}$ is PP_{cur} and g_j is a potential PP_{dest} of VR_j.

We use a linear weighted method to compute the total VRM cost. According to Equation 4.3 and 4.4, the total VRM cost of an individual, i, is represented as:

$$Cost(i) = \mathrm{w_{cost}} \cdot Cost\_a(i) + (1 - \mathrm{w_{cost}}) \cdot Cost\_b(i) \tag{4.5}$$

where $\mathrm{w_{cost}}$ is a weighted coefficient of two migration cost components.

Step 3: Fitness Assignment

When the evaluation step completes, the primary population and the secondary population are merged into a "temporary" population for fitness assignment. The size of this temporary population is the sum of P_{size} and S_{size} except at the first generation as the initial secondary population is empty. We apply a fitness assignment as in SPEA2 [136] on the basis of Pareto dominance. SPEA2 considers both the number of dominating and dominated solutions of each solution using two methods: dominance rank and dominance count.

Firstly, dominance rank records the number of candidate solutions by which a solution is dominated. A strength value is assigned to each solution according to the number of solutions it is dominated by. Let $S(i)$ be a strength value of a solution$i$. Hence, $S(i)$ is represented as:

$$S(i) = |\{\, j \mid j \in P_t + S_t \ \wedge\ i \succ j \,\}| \tag{4.5}$$

where P_t represents the primary population at the *t-th* generation and S_t is the secondary population at the *t-th* generation.+ stands for multi-set union so $P_t + S_t$ represents the temporary population. j is any individual in the temporary population. $|.|$ denotes the cardinality of a set which measures the number of elements in the set. $\succ$ is a Pareto dominance symbol which is defined in Section 2.3.5.1. $i \succ j$ represents that solution i dominates solution j. According to Equation 4.5, every solution in the temporary population has a strength value.

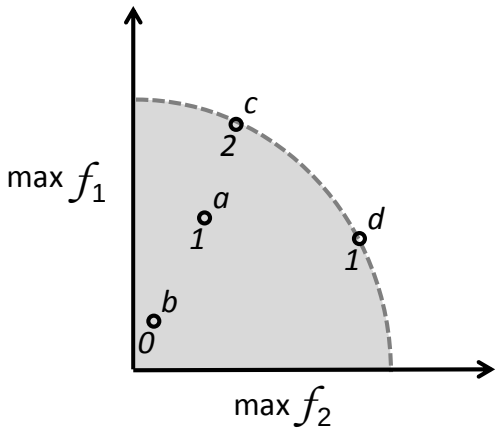

Figure 21. Example of Strength Values of Candidate Solutions

An example of strength values is shown in Figure 21. Assume that both functions need to be maximized and the Pareto front is represented by the dotted grey line. For points a, b, c and d, their strength values (labelled under the points) are determined by the number of solutions they dominate. For example, c dominates a and b, so its strength value is 2.

This example also describes why dominance rank information is not enough to determine the fitness value of solutions. Although both a and d have the same strength value, d is better than a because there is no solution that can dominate d whilst a is dominated by c. Hence, for a given solution, the number of solutions it is dominated by needs to be considered. This is considered in the dominance count method.

Then, after obtaining the strength value of solutions, a dominance count method is used for determining the fitness value. The fitness value of a certain solution is calculated by summing up the strength values of its dominators (solutions which dominate this solution). The fitness value, $R(i)$, of a solution a is represented as follows:

$$R(i) = \sum_{j \in P_t + S_t, j \succ i} S(j) \tag{4.6}$$

where j is any solution in the temporary population.

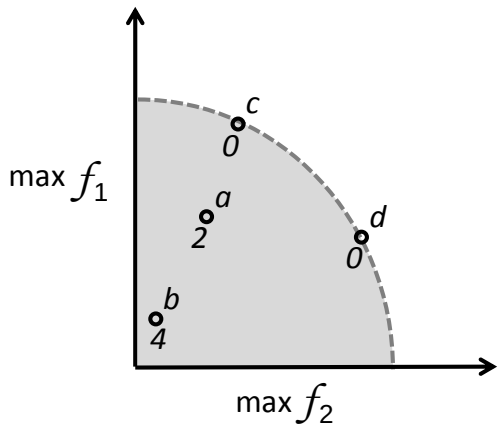

Figure 22. Example of Fitness Values of Solutions

An example of the fitness value of solutions is shown in Figure 22. For example, b is dominated by a, c and d, its fitness value is the sum of strength values of a, c and d (according to Figure 21). Based on the example, we can see a large fitness value means the solution is dominated by many solutions and conversely the fitness value of nondominated solutions is zero, e.g. c and d. Hence, it is important to note that a higher fitness value actually corresponds to the solution being less fit in SPEA2 fitness assignment. Furthermore, the fitness value of a solution may not be fixed. A solution (assuming it survives over several generations) can have different fitness values depending on its relationship with other solutions at each generation.

Step 4: Environmental Selection

Environmental selection determines which solutions in the primary and secondary populations (referred as the temporary population) form the next generation secondary population. We apply a truncation method [136] in environment selection to select a constant number of solutions to update the secondary population.

Firstly, all the nondominated solutions in the temporary population are copied into the next generation secondary population. The size of this nondominated solution set is denoted as T_{size}. Secondly, one of three actions is taken depending on the secondary population size S_{size} and T_{size}, namely:

- If $S_{size} = T_{size}$, the secondary population update completes.
- If $S_{size} < T_{size}$, the remaining places in the secondary population are occupied by the best dominated solutions from the temporary population, as necessary.
- If $S_{size} > T_{size}$, excess nondominated solutions are eliminated from the secondary population based on their density information given by the K-nearest measurements [146]. This means the solution with the minimum distance to another solution is iteratively eliminated until $S_{size} = T_{size}$. If two solutions have same distance with their nearest neighbours, the distance to the second nearest neighbour is compared. The solution with smaller distance to the second nearest neighbour is then eliminated. An example of this elimination process is given in Figure 23. Assume that two functions need to be maximized and dotted grey line stands for the Pareto front. If S_{size} is five, two solutions in Figure 23(b) of the seven potential solutions in Figure 23(a) are eliminated according to their distance to their nearest neighbours.

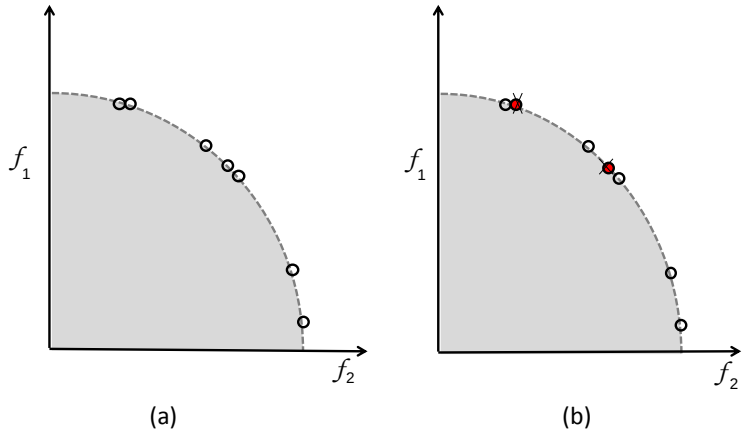

Figure 23. Example K-nearest Method

Step 5: Mating Selection

After the secondary population update completes in Step 4, mating selection is performed to determine which solutions are chosen from the secondary population to form the mating pool to reproduce offspring. The mating pool is cleared before the mating selection at every generation.

The selection mechanism we use is a tournament selection with replacement. It means that two solutions are randomly selected from the secondary population. After comparing their fitness, the better solution is copied into the mating pool. If both solutions have the same fitness value, one of them is randomly selected. The process repeats until the mating pool is full. It should be noted that no matter whether or not solutions are copied into the mating pool, they remain in the secondary population subject to the constraints (truncation method) given in Step 4. This is to ensure that good nondominated solutions persist between generations.

Step 6: Reproduction

In the reproduction stage, offspring is created by two operations: crossover and mutation, controlled by a crossover rate and a mutation rate. Before reproduction, the primary population is emptied ready for holding offspring.

Initially, two chromosomes from the mating pool are selected randomly to be the two parent chromosomes. Then, according to crossover and mutation operators, two children' chromosomes are generated. Next, two children are tested by the viability test (as explained in Section 4.4.3). The offspring that passes the test can be put in the primary population. Otherwise, the unqualified offspring is eliminated. Then, two chromosomes from the mating pool are selected randomly again to be two parent chromosomes. This process repeats until the primary population is full. Finally, increment generation counter $t = t + 1$ and go to Step 2.

Crossover and mutation operators are described in detail as follows.

Crossover produces offspring by exchanging gene information from the two parent chromosomes. The power of crossover is that it permits the combination of the best part of each parent to generate better offspring. In the binary representation, there are many kinds of crossover such as: one-point crossover, two-point crossover and uniform crossover. For real coded algorithms (such as integer and floating point representations), the performance of traditional crossover methods may be poor. Some real-coded crossover operators, such as arithmetical crossover, geometrical crossover and Blend crossover [149][150] have been proposed.

In this section, two crossover operator mechanisms that are used in the simulation are described. Their performance in the framework is shown in Section 5.6. Let $C_1 = (c_1^1, c_2^1, \dots c_n^1)$ and $C_2 = (c_1^2, c_2^2, \dots c_n^2)$be the two parent chromosomes which are selected to perform crossover and a child is denoted as H. Two crossover operators are described as follows.

1. Simple one-point crossover [159]

In simple one-point crossover, the position of crossover point is firstly randomly selected. Then, the two parent chromosomes are broken into two parts according to the crossover point. Finally, the first part of parent 1 and the second part of parent 2 are combined to be the first offspring. By analogy, the second offspring is generated by combining the second part of parent 1 and the first part of parent 2.

Let a randomly selected crossover point position be $\boldsymbol{l} \in \{1,2,\dots,n-1\}$ and two childrenbe H_1 and H_2. After one-point crossover, H_1 and H_2 are represented as:

$$H_1 = (c_1^1, c_2^1, \dots, c_l^1, c_{l+1}^2, \dots, c_n^2)$$

$$H_2 = (c_1^2, c_2^2, \dots, c_l^2, c_{l+1}^1, \dots, c_n^1)$$

An example is illustrated in Figure 24. It shows that the randomly selected position is 3. After crossover, the first and the second parts of the two parent chromosomes are swapped to form two child chromosomes.

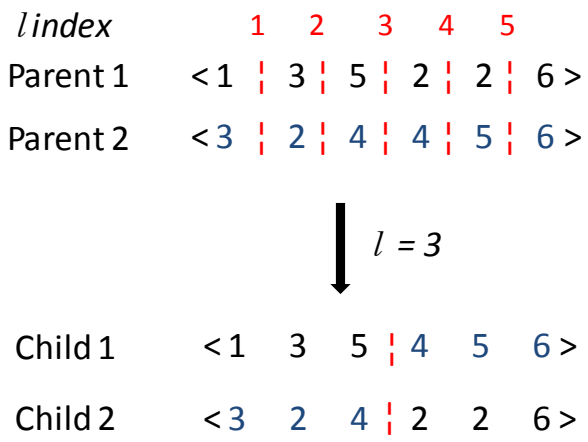

Figure 24. Example Simple One Point Crossover

2. Blend crossover

Blend crossover $(BLX-\alpha)$ is a popular real-coded crossover operators [151]. Let a child be $(h_1, h_2, \dots, h_n)$. The allele value of each gene h_i in a child is a randomly (uniformly) selected number of the interval $[H_i^1, H_i^2]$, where

$$c_{max} = \max(c_i^1, c_i^2)$$

$$c_{min} = \min(c_i^1, c_i^2)$$

$$I = c_{max} - c_{min}$$

$$\mathrm{H_i^1} = c_{min} - \alpha \cdot I$$

$$H_i^2 = c_{max} + \alpha \cdot I$$

andα is a positive parameter. Typically, α is equal to 0.5 [151].

The interval of random generation for a gene with a blend crossover operator is shown in Figure 25. An example of 6-node network is shown in Figure 26. Note that the values in the interval $[H_i^1, H_i^2]$ are not all feasible. Take a 6-node network as an example; the feasible allele values are from 1 to 6. However, according to crossover definition, the selected interval can exceed the feasible interval ([1, 6]). If this happens, the allele value of the gene is selected from the feasible interval.

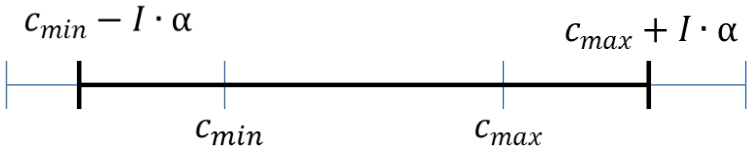

Figure 25. Interval of Random Generation for a Gene with Blend Crossover

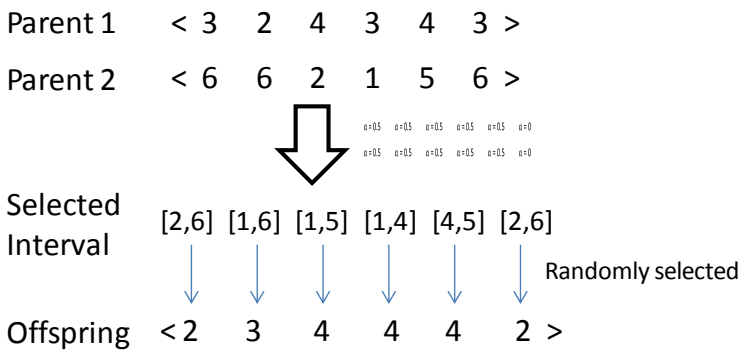

Figure 26. Example Blend Crossover

Mutation is used for maintaining the genetic diversity from one generation to another by altering some gene values randomly. For each gene in a chromosome, a uniform random variable is firstly generated in the interval [0, 1]. If the random variable is smaller than the user-defined mutation rate, the gene can be modified to any feasible value by a random method. Mutation operators may alter one or more gene values in a chromosome. An example of mutation operator is shown in Figure 27.

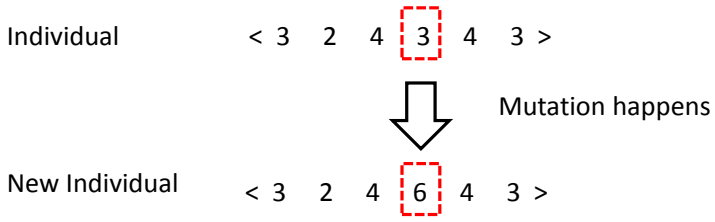

Figure 27. Example Mutation Operator on the 4thGene

Step 7: Final Solutions

When a terminal condition, which is the maximum number of generations in our case, is reached, the simulation stops. Then, the secondary population in the final generation is the outcome.

4.5 **Summary**

An important consideration in the dynamic energy management framework is how to select appropriate physical platforms to move virtual router instances to, given various constraints. Chapter 4 describes a novel evolutionary-based algorithm called VRM_MOEA to solve the problem.

Firstly, this problem is described in Section 4.2and some conditions are required to be satisfied in a potential algorithm in Section 4.3. Then, an evolutionary-based algorithm called VRM_MOEA is developed to solve the problem in Section 4.4. We use an integer encoding to represent an individual since PPs and VRs are labelled with decimal numbers. In order to make sure every individual satisfies the migration constraints, a viability test is used for filtering out the unqualified solutions.

Finally, the steps of VRM_MOEA are described in detail in Section 4.4.4. There are three populations in the algorithm, i.e. the primary population, the secondary population and the mating pool. The primary population is a regular population that is replaced by offspring every generation. The secondary population is used to archive nondominated solutions so no matter how many generations are iterated, the best nondominated

solutions are retained. The mating pool is used for accommodating parent chromosomes that are used to produces offspring. By selection and reproduction mechanisms, these three populations are updated at each generation until the terminal condition is reached, i.e. the maximum number of generations in our case. The output of the algorithm is a group of solutions, which are stored in the secondary population.

CHAPTER 5 SIMULATION MODELLING AND PARAMETER SETTING

5.1 Introduction

A dynamic energy management framework is proposed in Chapter 3 and an evolutionary-based algorithm VRM_MOEA is developed to address the destination physical platform selection problem in Chapter 4. The next step is to build a valid, credible and appropriately detailed simulation model in order to evaluate the performance of the framework. Hence, a stochastic and discrete-event simulation tool was created and is described in this chapter.

The remainder of Chapter 5 is organized as follows. Initially, various components of the simulation tool are introduced in Section 5.2, including simulation entities, the components of discrete-event driven simulation, the simulation events, the event graph and the simulation flow chart. Then, several functional simulation modules are described from Section 5.3 to Section 5.5, e.g. the network topology and traffic models. Next, we explore the appropriate parameter values to be used in VRM_MOEA, e.g. the mutation rate, crossover rate and the number of offspring in Section 5.6. Next, the validation of simulation model is described in Section 5.7. Finally, Section 5.8 presents a chapter summary.

5.2 Simulation Tool

There are various means by which the framework could be simulated including the use of existing software (e.g. some commercial software or open source programs) or

building a new bespoke simulator. Generally, existing software can provide an extensive library of models and sometimes it is easier to debug. However, as our framework is new, no suitable protocols or architectural models are present in existing software; these would have to be created. Furthermore, we also need a means of simulating packet flows over high-speed optical links over many hours. This is not feasible with a traditional packet-level simulator. Instead we constructed a new hybrid fluid flow / packet simulator that can achieve this and possess all the features we required.

This simulation tool was built using C/C++ with Microsoft Visual Studio 2012 and enabled us to evaluate the performance of the dynamic energy management framework over many hours of simulated time. The data traffic is modelled as time-varying flows and signalling information is modelled as discrete packets. The optical channel resources are also modelled as finite countable entities (e.g. the number of wavelengths per channel) and the wavelength continuity constraint is applied when new optical connections are constructed in response to VRM events, i.e. establishing optical connections to remotely located VR instances.

The simulation tool was also created based on a stochastic and discrete-event driven simulation. Firstly, a stochastic simulation means that there is at least one random input component in the system which results in a random output [155]. Thus, the output of a simulation trial (or run) is regarded as an estimate of the true characteristics of the model. In order to get close estimate of the true characteristics, multiple simulation runs with different random number seeds are required. Secondly, a discrete-event simulation indicates that system operations are composed of a series of events. An event is defined as an instantaneous occurrence which changes the state of the system [154]. The order of events is controlled by a simulation clock and chronological sequencing of the event-list. When the leading event is removed from the event-list, the simulation clock moves to the time stamp of this event and the period between the old event and the new event is skipped. As the discrete-event simulation is driven by a sequence of events instead of fixed time increments, it can greatly accelerate the simulation.

5.2.1 **Simulation Entities**

Simulation entities are objects, which request resources or communicate with other entities in a simulation. Simulation entities and a brief description are shown in Table 1.

Table 1. Simulation Entities

Entity Name	**Description**
Source	An object that generates the traffic. The traffic contains information such as source and destination.
Destination	The final object that receives the traffic.
Physical Platform	An object that hosts one or more VR instances. It has finite resource such as switch (fabric) capacity and the number of line-cards.
Virtual Router	An object that processes and forwards the traffic.
ROADM	An object that adds, drops and switches optical signals in the optical domain.
Fiber	An object that contains a group of finite optical channel resources.
Channel	An object which is used for transmitting the traffic in the optical domain
Centralized Control Unit	An management unit which collects the network condition information, computes VRM_MOEA, coordinates VRM events, etc.
Population	An object that holds a group of candidate solutions in VRM_MOEA. It can be modified to be a primary population, a secondary population and a mating pool.
Individual	An object which contains the mapping information of VRs and PPs.

5.2.2 **Components of Discrete-Event Driven Simulation**

A discrete-event driven simulation is used to create an imitation of a system with a computer program. Computer programs generally have a clear structure. Likewise,

discrete-event driven simulations use some common components [155]. The components of our simulation are described as follows.

Simulation clock: a variable which presents the current value of simulation time. When the earliest (leading) event is extracted from the event list, the simulation clock moves to the time stamp of this event and the period between the event and the previous executed event is skipped. As an event occurs instantaneously, no simulation time elapses whilst the event occurs. Note that the unit of the simulation clock can be any suitable period of time in the real world. In our simulation, the unit of the simulation clock is one second.

System state: a group of variables which describe the system at a specific moment. At least one of these state variables changes when an event occurs. An example of system state in our simulation is the state of a PP (active or sleep). If a PP does not host any VR instances, the state of PP can be changed to "sleep".

Event list: a "to do" list that controls the timing of the occurrence of events. The event list is effectively an event schedule, which ranks the events based on their occurrence time. The event list may be composed of many entries and each entry contains the event information, e.g. event time, event type and some variables need to be the input of an event. Typically, events are added or removed dynamically during the simulation. An event is usually generated by another event and then is inserted in the list according to its intended event occurrence time. On the other hand, an event is eliminated from the event list when the event is complete. The types of events in our simulation are explained in Section 5.2.3.

Statistics: a set of variables that are used for storing the statistical information of interest to track the performance of the system. For instance, the accumulated energy consumption of the network and the number of occupied optical channels in our case are tracked during the simulation.

Initialization routine: a subprogram which initializes the simulation model, e.g. simulation clock, system state and statistics and initial events. The initialization routine is invoked when the simulation clock is equal to zero.

Timing routine: a subprogram that obtains the next event from the event list and the simulation clock moves to the time stamp of next event.

Event routine: a subprogram that updates the system state and statistics when an event occurs. Note that each type of event has its own event routine. Each event routine is described in detail in Appendix B, Part 1.

Random number generator: a subprogram or a set of subprograms to generate random numbers. The random number generator in the simulation is introduced in Section 5.5.

Terminal condition: a predefined condition to end the simulation. Generally, the triggering of events may result in the creation of other events. Hence, a discrete event simulation theoretically can run indefinitely. Thus, a terminal condition is required to stop the simulation. Typical conditions include that setting the maximum number of events or defining a maximum value of a statistical measure. In our case, the maximum simulation clock time is used because we are interested in the network operations over a specific period of time (e.g. one day or one week).

Report generation: a subprogram which collects and computes system statistics. When the terminal condition is achieved, a report is generated as the output of the simulation. In our simulation, the report contains information such as the results of VRM_MOEA algorithm, the total energy saving and the time of VRM events.

The relationships between these components are described in the simulation flow chart in Section 5.2.5.

5.2.3 **Simulation Events**

The basic element of a discrete-event driven simulation is an event, which is defined as an instantaneous occurrence that changes the system state. In this section, we identify

nine types of events based on the network operations in Table 2. The event routine of a particular event is described in detail in appendix B, Part 1. The logical relationship amongst events is described in Section 5.2.4.

Table 2. Types of Events

Event Type	Event Description
Src_gen	The change in flow rate is generated at a source.
Arrival	The traffic arrives at a VR instance.
Departure	The traffic leaves a VR instance.
Reach_destination	The traffic reaches a destination.
Collect	The CCU collects the network condition information.
VRM_MOEA	The VRM_MOEA algorithm is computed.
Optical_test	Used for examining whether a candidate solution can be supported by the optical resources.
VRM	Virtual router instance(s) move to their destination physical platform(s).
End_simulation	End of simulation.

5.2.4 **Event Graph**

In this section, we use a simplified event graph to present the relationships amongst different types of event. An event graph [156] is a popular method of graphically representing the logical relationships within discrete-event driven simulations. In the event graph, an event is represented by a node (vertex) and nodes are connected by arcs (edges). Arcs indicate the relationship between events. The event graph of our network system is shown in Figure 28.

There are several notations to indicate types of events in Figure 28. An event, which is at the end of a thin jagged arrow, is required to be scheduled when the simulation clock is zero as the event cannot be scheduled by other events. In our case, *Src_gen, Collect* and *End_simulation* events need to be scheduled initially. Meanwhile, an event, which is at

the end of a heavy, smooth straight arrow or a heavy curved arrow, can be scheduled by the event at the beginning of arrow after a time delay. For example, the *Src_gen* event can re-schedule itself and schedule an *Arrival* event.

In addition, a label above a heavy straight arrow represents the condition needed to schedule the following event. If the condition is not satisfied, the following event is not scheduled. In Figure 28, the label "C1" represents that one of the migration conditions needs to be satisfied so a *VRM_MOEA* event can be scheduled. Similarly, "C2" means when a candidate solution is supported by the optical resource, the following *VRM* event can be scheduled.

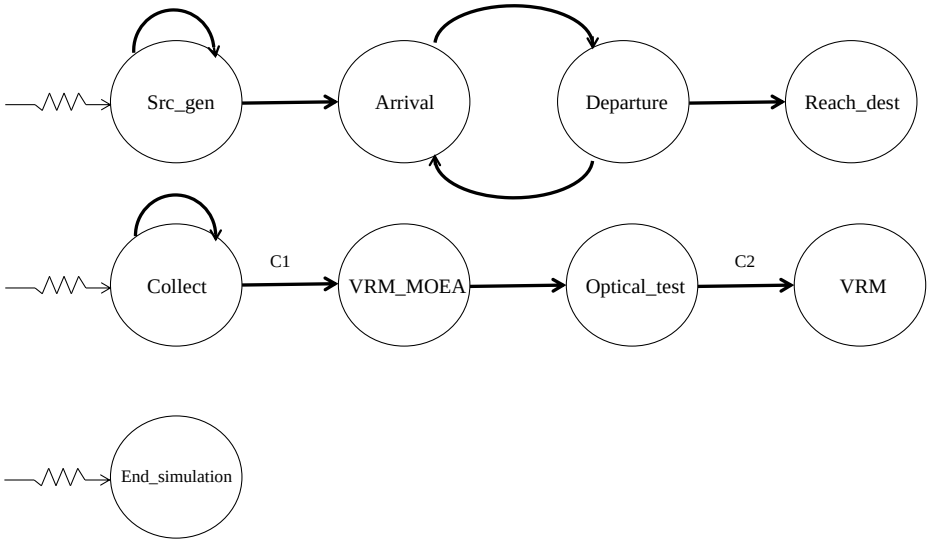

Figure 28. Simulation Event Graph

5.2.5 **Simulation Flow Chart**

The flow chart of the implemented discrete-event driven simulation is shown in Figure 29. An initialization routine is firstly invoked. The initialization routine includes setting the simulation clock to 0, reading parameters and constants from input files (e.g. a network topology configuration file) to initialize the network architecture, and initializing the system states and statistics. An event list is also created which contains *Src_gen, Collect* and *End_simulation* events which cannot be scheduled by other events.

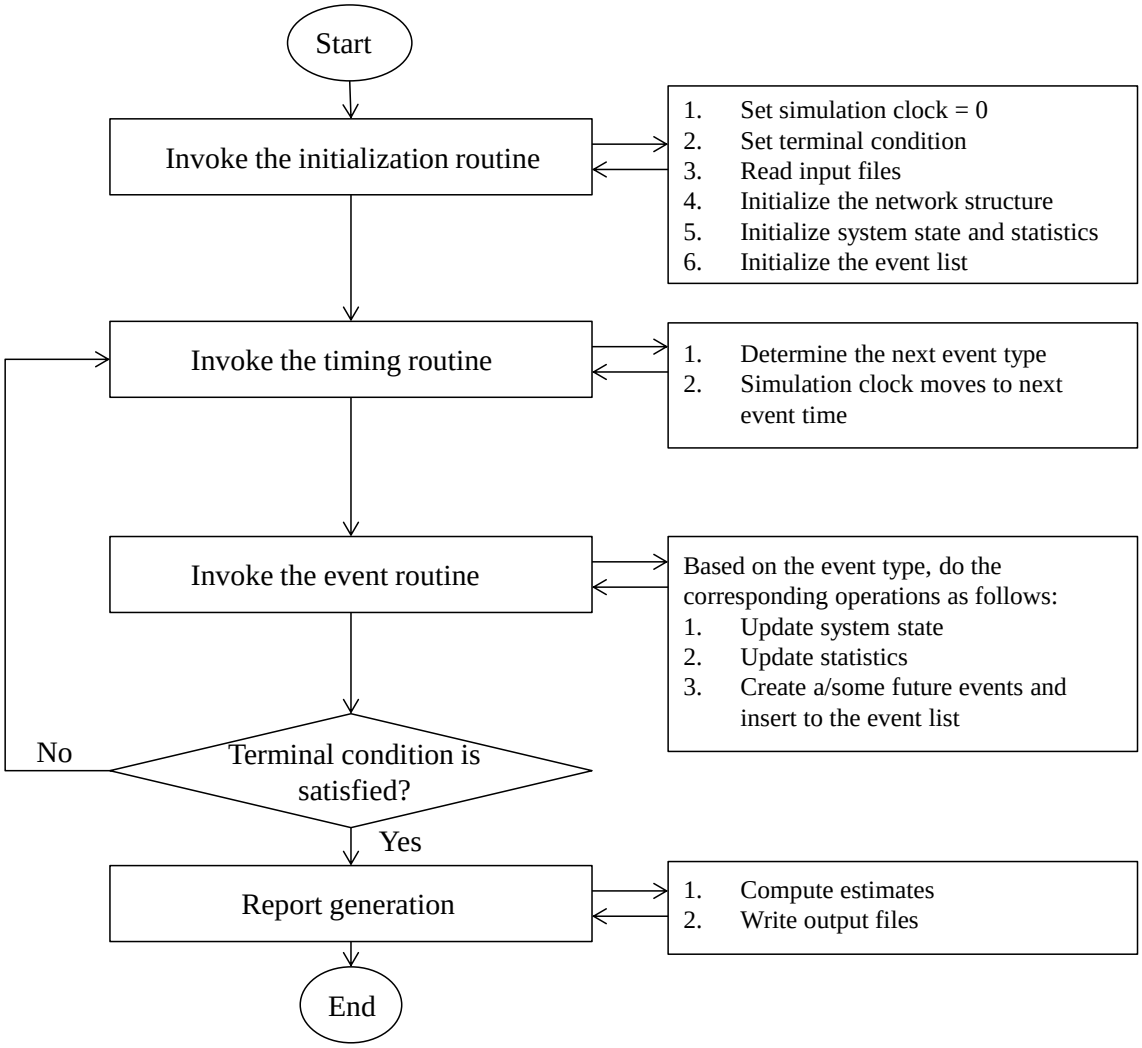

Figure 29. Simulation Flow Chart

When the initialization routine completes, the timing routine is then invoked to fetch the next event and the simulation clock is advanced to this event (simulated) time. According to the type of next event, the corresponding event subroutine is invoked. The event subroutine usually includes updating the network state and statistics, scheduling further events and inserting them into the event list.

Then, the simulation terminal condition is examined. If the terminal condition is satisfied, the estimates are computed and then exported to output files. The simulation finally ends. Otherwise, the simulation continues, i.e. the timing routine is invoked again

to fetch the next event. In our case, the maximum value of simulated time is used as the terminal condition, which is contained in an *End_simulation* event. Hence, when the *End_simulation event* is fetched, a report is generated and the simulation then stops.

5.3 Network Topology

In this section, two types of network topology are introduced, i.e. three reference network topologies from previous work and a simple network topology generator that is used for investigating the relationship between the network node degree and the framework performance.

5.3.1 Reference Network Topology

Three reference networks are used, i.e. a 6-node-8-link small size network (6N8L) [27], the National Science Foundation NETwork (NSFNET) with 14 nodes and 21 links (14N21L) [27] and the Abilene network with 12 nodes and 15 links (12N15L) [169]. The numbers along the arcs are the span lengths between nodes and unit is kilometers. The topologies of these networks are shown from Figure 30 to Figure 32.

5.3.2 Random Network Topology Generator

One factor, which may affect the performance of the dynamic energy management framework, is the network topology. An important metric of the network topology is the average network node degree. In our scenario, the average node degree represents the ratio of the number of PPs to links.

The network node degree may affect the performance of the framework. For example, a VR instance has more choice of destination PPs in a network with a higher node degree if the same migration cost is considered. Hence, besides the three reference network topologies described in Section 5.3.1, we also developed a simple Random Network

Topology Generator (RNTG) for exploring the relationship between the performance of framework and the average network node degree.

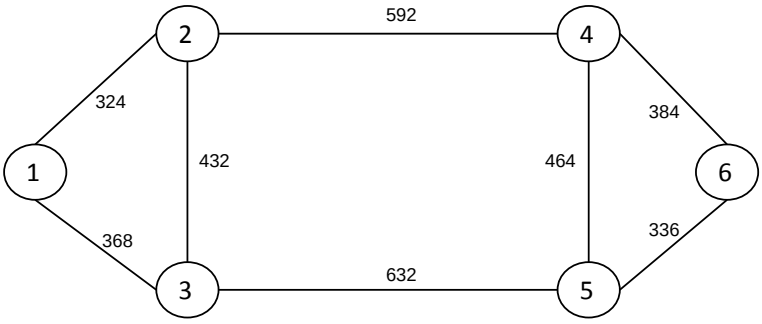

Figure 30. 6-node-8-link Network Topology

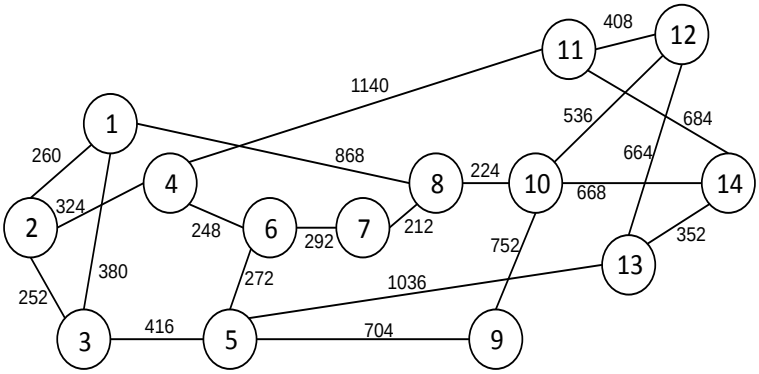

Figure 31. NSFNET Network Topology

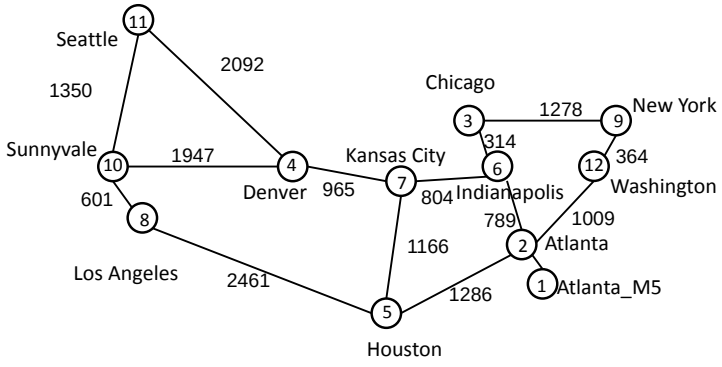

Figure 32. Network Topology of Core Abilene Network

Network Topology Generators (NTGs) are used typically to generate network topologies using some real Internet characteristics to test the performance of network protocols in the simulation. In theory, the network protocols are designed to be independent of the underlying network topology. However, network topologies can impact the performance of network protocols [170]. In order to test the performance of protocols, realistic network topologies need to be generated.

How to generate network topologies with real Internet characteristics is an open question. Several generations of NTGs have been proposed. The first generation NTGs were developed based on the assumption that a network topology is totally random. A typical such NTG is Waxman [171], which generates uniform nodes in a plane and the link creation probabilities are determined by the Euclidean distance between the nodes. Later research revealed that real network topologies are not random, but they have a deliberately hierarchical structure [172]. Representative NTGs are Transit-Stub [173] and Tiers [174] in this generation. Then, the authors of [175] indicated the degree distribution to follow a power-law in the Internet. This discovery makes the beginning of the third generation of NTG [176]-[179].

However, the NTGs above are usually developed for generating the large-scale structure of the Internet, which can consists of thousands of nodes. In our study, the objective networks are the Tier-2 networks which usually have dozens of nodes, such as 14-node-21-link NSFNET and 12-node-15-link core Abilene network. Thus, these existing NTGs may not suitable for our study and we developed a simple RNTG instead.

The RNTG generates the network topologies randomly for a given number of nodes and links. All links are bidirectional. The main concept is to first connect two nodes together. A third node is then connected to a randomly selected node from the first two. The remaining nodes are linked with a randomly selected node from the previously introduced nodes. After all nodes are linked into the network, the remaining links are used for connecting any two randomly chosen nodes. The algorithm is described in Table 3 and the RNTG verification is shown in Appendix B, Part 4.

Table 3. RNTG Algorithm

Algorithm: RNTG

Input: NUM_NODE; //the number of nodes in the network

Input: NUM_LINK; //the number of links in the network

Data: i; //counter

Data: $Node_i$; //the *i* th node in the network

Data: Node_A, Node_B; // any node in the network

Output: Degree [NUM_NODE]; // array of degree number of each node

Output: Connection_Matrix [NUM_NODE][NUM_NODE-1]; // network connection of each node

Begin

1 Connect(Node1, Node2)

2 For i = 3 to NUM_NODE

3 Choose a Node_A randomly from the existing network

4 Connect (Node_A, $Node_i$) , Node i is added into the network

5 End for

6 For i = NUM_NODE -1 to NUM_LINK

7 Choose any Node_A and Node_B satisfying Node_A!=Node_B and no link between Node_A and Node_B

8 Connect(Node_A, Node_B)

9 End for

End

5.4 **Traffic Model**

The data traffic is modelled as time-varying flows. Although data traffic is stochastic in nature giving rise to significant variation about the average flow rate, in our case, we are dealing with aggregated flows over relatively long timeframe. Under these circumstances, it is common to represent the traffic using a fluid approximation [161] [162]. In the simulation, we used two types of traffic, i.e. a sinusoidal traffic model and an historical traffic demand matrix taken from the Abilene network.

Typically, the traffic in a Tier-2/3 network has a regular diurnal pattern based on people's activities, which is high during working hours and much lighter in hours

associated with sleep [40] [41]. In order to explore the effect of traffic variation with time, a sinusoidal function is used for modelling the traffic with a daily periodicity [30]. We also assume that for each pair of peering nodes in the network, they have a similar daily traffic pattern.

Given an N-node network, let T_{ij} be the traffic demand from a source node $i = 1,2, \dots, N$ to a destination node $j = 1,2, \dots, N$, then T_{ij} at simulated time t is:

$$T_{ij}(t) = t_{ij} \cdot \left[\frac{1-\gamma}{2} \cdot (1 + \sin(ft)) + \gamma\right] + \sigma \quad (5.1)$$

where t_{ij} is the amount of traffic going from node i to j and σ is a random noise variable. In addition, γ is a parameter to control the percentage of the peak traffic that is used to represent the off-peak traffic. For example, if the amount of off-peak traffic is equal to 20% of the peak traffic, γ is equal to 0.2. Meanwhile, f is a parameter related to the simulation time unit. For instance, if the unit is 1 minute in the real world, then $= \frac{2\pi}{1440}$.

We also used a historical traffic demand dataset in order to investigate the performance of framework over the real network data. In research community, it is not difficult to find a dataset of single or several capture traffic traces [182] [183] or BGP routing data [184] [185]. However, the traffic demand matrix of a network is rare. This is because for ISPs, the traffic demand information is usually protected and not published for security reasons. Till now, as far as we know, the available publicly network traffic dataset are from the GEANT [40] and the Abilene network [186].

In our study, we used the Abilene network dataset. The Abilene network is a backbone network created by the Internet2 community. The clients of Abilene network are mostly universities and affiliate institutions in the United States [187]. The core network has 12 nodes and 15 links. The traffic demands of 144 pairs of sources and destination are collected every five minutes over a discontinuous 24 weeks. The unit of traffic is kbps. The traffic demand over the 1stweek, the 13th week and the 24th are shown in Appendix B, Part 2.

5.5 **Random Number Generator**

The random number generator we used in the simulation is a combined Multiple Recursive Generator (MRG) [155], which supports up to 10,000 streams with seed vectors spaced 10^{16} apart. Details are given in Appendix B, Part 3.

5.6 **VRM_MOEA Parameter Setting**

An essential issue in MOEAs is to determine appropriate parameter values, e.g. the mutation and crossover rate and the number of offspring. This is because parameter values usually have influence on the performance of a MOEA, i.e. whether the algorithm can find a Pareto set approximation and whether it can find such a group of solutions efficiently.

Traditionally, these parameter values can be selected based on some conventions. For example, the mutation rate is usually low (e.g. 0.01 or 0.1) and crossover rate is usually high (e.g. 0.95) [188]. However, because a given application usually has its own features, these conventional values may not always be good. Under these circumstances, experimental comparisons are usually required in order to determine the appropriate parameter values.

Generally, determining appropriate parameter values (i.e. parameter setting) in MOEAs is not a trivial problem; sometimes it is a time-consuming task. Much effort has been spent to develop good heuristics for parameter setting. These parameter setting methods are commonly divided into two types: parameter tuning and parameter control [189]. In the case of parameter tuning, the aim is to find good parameter values before running the algorithm. During the run, each parameter is configured with a fixed value. Alternatively, the parameter control method starts a run with initial parameter values and suitable control strategies; and then the parameter values are modified automatically based on control strategies during the run-time of the algorithm. Compared with

parameter tuning, parameter control is more complex. Nevertheless, for a specific application, even parameter tuning can be difficult because of limited knowledge concerning the relationship between parameters and the performance of the algorithm as well as the large number of parameter options [190]. For example, it is not easy to select a good mutation rate $p_m \in [0,1]$ as its domain contains infinite options. One solution is to run the experiments with several sample values $(0, 0.1, \ldots, 1.0)$ and then to select a good mutation rate value based on the algorithm performance.

Given this background, we explore how to select appropriate parameter values of VRM_MOEA. The remainder of the section is organized as follows. Initially, Section 5.6.1 introduces the objective of parameter setting and two measurement metrics. Then, Section 5.6.2 describes the impact of various parameters on the performance of VRM_MOEA with associated discussions. Finally, some appropriate parameter values are summarized in Section 5.6.3.

5.6.1 **Objective and Performance Metrics**

In our study, we use a parameter tuning method to find the appropriate parameter values based on experimental comparisons. There is a two-fold objective for tuning the parameter values in VRM_MOEA as follows:

1. *Complete VRM_MOEA in a limited period of time*.
2. *Obtain a group of reasonable solutions.*

In summary, the objective is to find a group of good solutions efficiently in a limited time. Typically, the network state information is collected every 15 minutes to examine whether the network conditions satisfy the need for VRM. If this is the case, VRM_MOEA is computed. Hence, VRM_MOEA is an online algorithm that needs to obtain a group of reasonable solutions quickly.

In order to measure the performance of VRM_MOEA with various parameter values, two measurement metrics are used, i.e. the computational effort and the quality of solutions.

Firstly, there are several ways to measure the computational effort of MOEAs, e.g. the number of fitness evaluations and the CPU computational time. As we focus on completing the algorithm in a limited period of time, we use CPU computational time to represent the computational effort. Secondly, comparisons of different simulations with various parameters are not easy in MOOPs. Hence, several performance metrics have been proposed [152] [153]. We selected a popular performance metric called *Hypervolume* [152] to measure the quality of solutions. *Hypervolume* is a performance metric considering the size of the objective space covered by the solutions. A higher value of *Hypervolume* indicates a group of solutions possess better quality as they cover the bigger share of the objective space.

In order to calculate the size of the objective space covered by the solutions, a point called the reference point is needed. Hence, setting a suitable reference is important for calculating the value of the *Hypervolume*. Generally, the reference point $\boldsymbol{c}$ is set to be the original point in maximization problems, e.g. (0,0) in a two dimensional objective space. On the other hand, in minimization problems, $\boldsymbol{c}$is usually set to exceed the maximum values of each objective.

Formally, let $\boldsymbol{A} = \{\boldsymbol{a_1}, \boldsymbol{a_2}, \dots, \boldsymbol{a_n}\} \subseteq \boldsymbol{X}$ be a set of n decision vectors. The function $\boldsymbol{M_H(A)}$ gives the *Hypervolume* enclosed by the union of the polytopes $\boldsymbol{p_1}, \boldsymbol{p_2}, \dots, \boldsymbol{p_n}$, where each $\boldsymbol{p_i}$ is formed by the intersections of the following hyperplanes arising out of $\boldsymbol{a_i}$, along with the axes: for each axis in the objective space, there exists a hyperplane perpendicular to the axis and passing through the point $(f_1(a_i), f_2(a_i), \dots, f_k(a_i),)$[152].

For example, in a two dimensional (2-D) objective space, each $\boldsymbol{p_i}$ of a solution $\boldsymbol{a} \subseteq \boldsymbol{X}$ represents a rectangle defined by the reference point $\boldsymbol{c} = \left(c_{f_1}, c_{f_2}\right)$ and $(f_1(a), f_2(a))$. The size of the union of all such rectangles covered by the solutions is used as the measure:

$$\boldsymbol{M_H(A)} = \boldsymbol{\lambda}\left(\cup_{\boldsymbol{a}\in\boldsymbol{A}}\{\,[\boldsymbol{f_1(a)}, c_{f_1}] \times [\boldsymbol{f_2(a)}, c_{f_2}]\,\}\right) \qquad (\mathbf{5.2})$$

where $\boldsymbol{\lambda}(\cdot)$is the standard Lebesgue measure.

The graphical representation of *Hypervolume* in a 2-D maximization problem and a 2-D minimization problem are shown in Figure 33(a) and Figure 33(b), respectively. The blue dots stand for the solutions in the objective space and the black dot is the reference point. The grey area represents the *Hypervolume* which is the union of the rectangles that are defined by the solutions and the reference point.

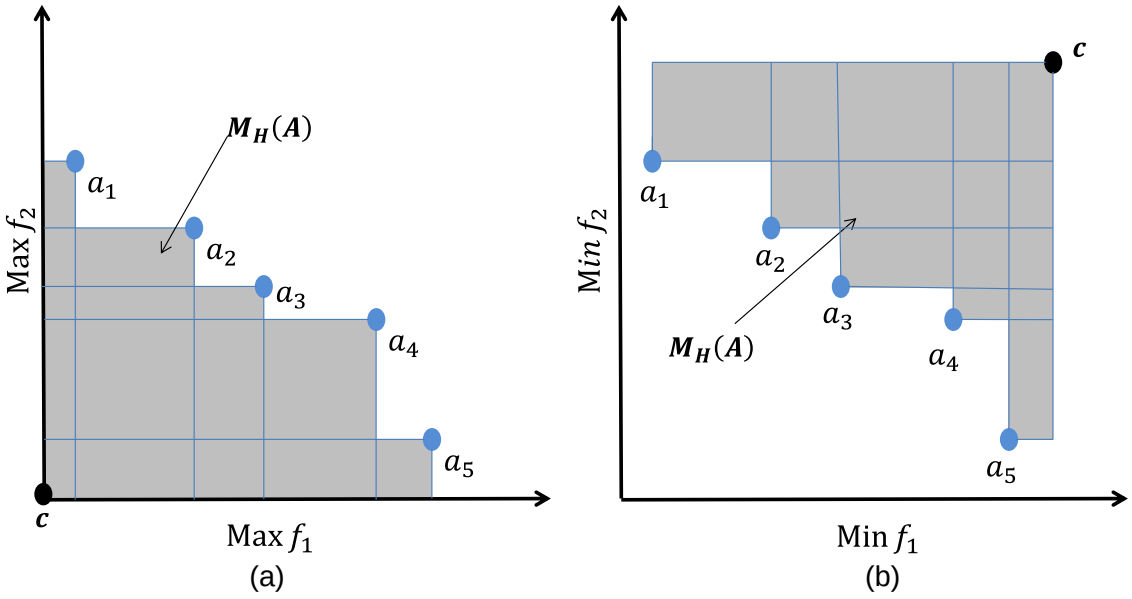

Figure 33. Graphical Representation of *Hypervolume*

5.6.2 Parameter Impact

In this section, the impact of different parameters in VRM_MOEA is investigated. Firstly, the appropriate values of mutation and crossover rate are explored in Section 5.6.2.1. Then, the impact of the maximum number of generations, the number of offspring and mating pool size are described from Section 5.6.2.2 to Section 5.6.2.3, respectively. Next, the impact of traffic load and number of nodes are investigated in Section 5.6.2.5 and Section .5.6.2.6.

5.6.2.1 Mutation and Crossover

In this section, we explore the appropriate values of mutation and crossover rate in VRM_MOEA. In VRM_MOEA, the mutation rate is the probability of a mutation happening per gene per generation. Hence, the range of the mutation rate is [0, 1].For example, if the mutation rate is 0.5, there is 50% chance that the mutation occurs on a gene per generation. Similarly, the crossover rate indicates the probability that two parent chromosomes experience a crossover. The range of the crossover rate is [0, 1].

In order to measure the performance of VRM_MOEA with various mutation and crossover rate values, two measurement metrics are used as explained in Section 5.6.1, i.e. the computational time and the quality of solutions.

Note that from Figure 34 to Figure 36, the step interval of the mutation rate and crossover rate is 0.1 and the maximum number of generations is 5000. The mutation mechanism used is "point mutation" and the crossover mechanism is "one point crossover" (as described in Section 4.4.4, Step 6). For the sake of simplicity, the average values of 10 simulation runs with random seeds are shown.

Firstly, some experiments are conducted on the computational time with different settings of mutation and crossover rate as shown in Figure 34.

It is clear that for a certain crossover rate, the computational time increases with the mutation rate. This is because a high mutation rate leads to a random search which increases the probability of failure in the viability test. That is to say, with increasing mutation rate, more genes are mutated in a chromosome and the offspring may be effectively generated randomly. As the viability test (as described in Section 4.4.3) is used to examine if the offspring satisfies the migration constraints, the randomly generated offspring may not be able to pass the test. If an offspring cannot pass the test, it is eliminated from the offspring population and then the reproduction procedure (including the mutation mechanism) repeats until the required number of viable offspring is achieved.

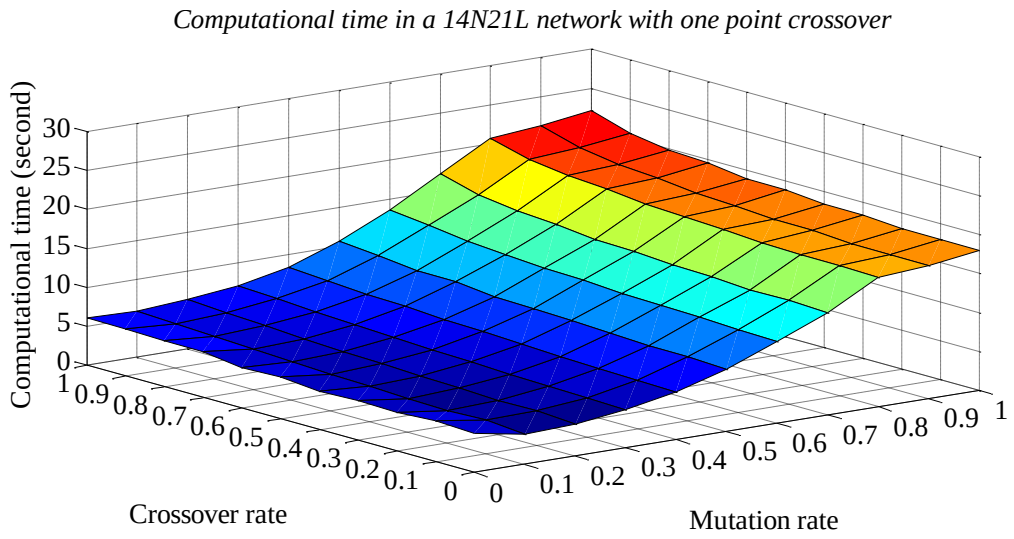

Figure 34. Computational Time of Various Mutation and Crossover Rates

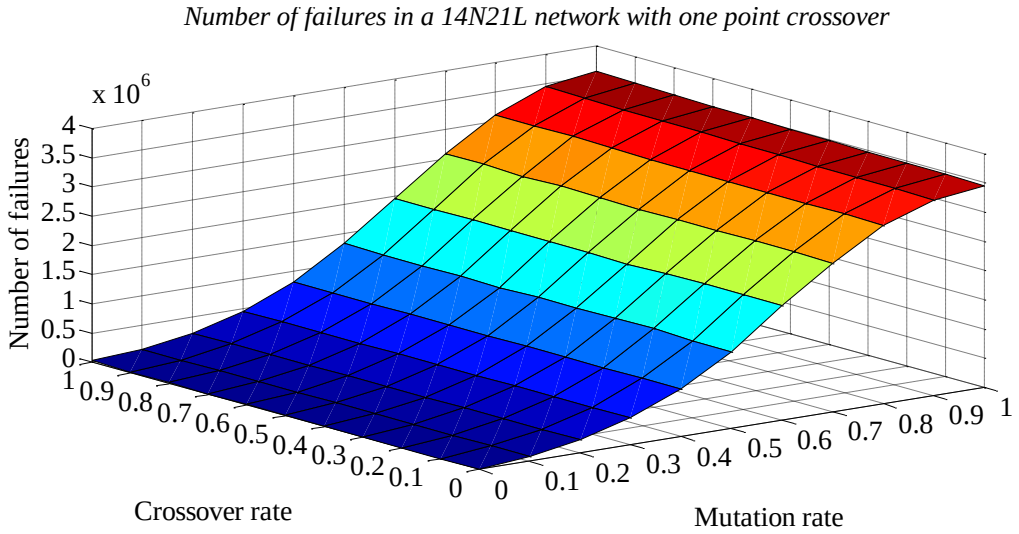

Figure 35. The Accumulated Number of Failures of the Viability Test with Various Mutation and Crossover Rates

In order to observe the effect of the viability test, the accumulated number of failures of the viability test is explored with various mutation and crossover rates. Figure 35 shows the accumulated number of failures when the maximum generation count is set to 5000. It is clear the number of failures has a similar trend to the computational time in Figure 34. It indicates that with a higher mutation rate, more individuals need to be

generated in order to have a certain number of qualified individuals (offspring that passes the viability test), which result in a longer computational time.

Furthermore, when the mutation rate is from 0 to 0.3 and from 0.8 to 1, the increasing rate of the computational time is smaller than the rate in the range from 0.3 to 0.8. This means that changes of mutation rate have a larger impact in the range [0.3, 0.8]. On the other hand, for a certain mutation rate, the computational time of the cross rate has little variation. The crossover mechanism has only a small impact on the computational time.

Secondly, the impact of different mutation and crossover rates on the quality of solutions is shown in Figure 36 using the *Hypervolume* metric. Based on the *Hypervolume* definition in Section 5.6.1, a population with better quality has a higher value of *Hypervolume*. In Figure 36, we compare the values of *Hypervolume* when the maximum generation count is set to 5000.

On the one hand, for a certain crossover rate, the values of *Hypervolume* are lowest when the mutation rate is 0 (no mutation); when mutation rate is from 0.1 to 0.4, the values of Hypervolume are highest. This means that configuring the algorithm with a small mutation rate has good performance. Then, with an increasing mutation rate from 0.4 to 1, the values of *Hypervolume* quickly reduce. This is because a high mutation rate leads to a random search which does not help to improve the quality of solutions. Thus, in this case a small mutation rate [0.1, 0.4] is preferred.

Nevertheless, for a certain mutation rate, the values of *Hypervolume* have little variation with various crossover rates. The crossover mechanism has little impact on the performance of VRM_MOEA. This result is similar with the computational time result shown in Figure 34.

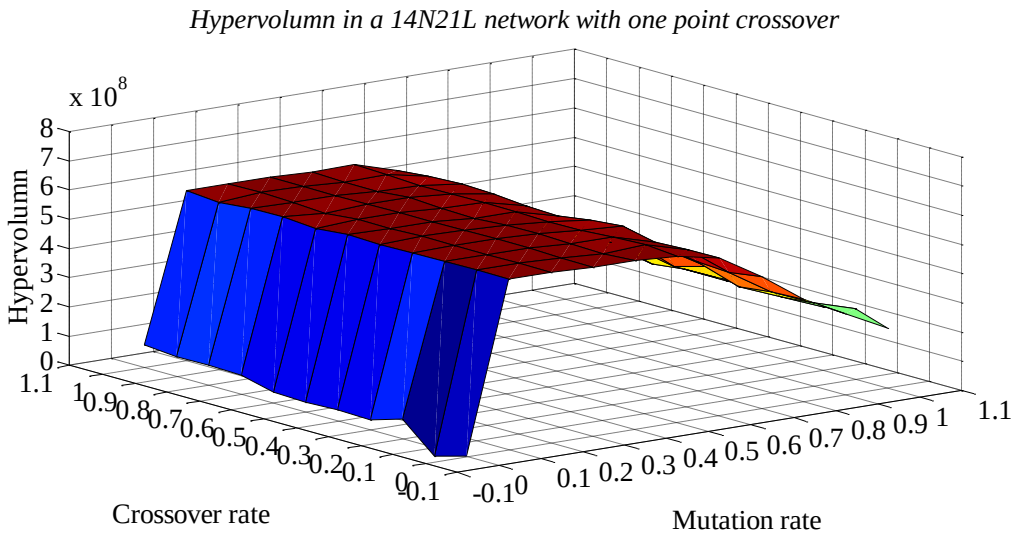

Figure 36. *Hypervolume* with Various Mutation and Crossover Rates

Figure 36 indicates that the mutation mechanism is necessary and important within VRM_MOEA as the algorithm performance is the worst when no mutation is performed. In contrast, without one point crossover, VRM_MOEA still shows good performance with a low mutation rate. Two possible explanations are as follows:

1. It indicates that the single-point crossover may not be an appropriate crossover mechanism in VRM_MOEA as it seems do not help to improve the quality of the solutions. More effort may be needed to search for an appropriate crossover mechanism. Hence, another popular crossover mechanism, called BLX-a (as explained in Section 4.4.4), is used in the experiments (the results are shown in Appendix C, Part 1). The trend in the results of computational time and the quality of solutions with BLX-a crossover are similar to that of the single-point crossover. This leads to the second explanation.

2. According to the results of two crossover mechanisms, the crossover mechanism may have limited impact on the performance of VRM_MOEA. This may come from the nature of destination physical platform selection problem. Swapping parts of the parent chromosomes usually does not bring benefit and sometimes even produce offspring of worse quality. In order to describe it more clearly, an example is shown in Figure 37.

Assume that there are two candidate solutions that try to consolidate VR instances onto fewer PPs to save energy. These two solutions are good in the population and selected as parent chromosomes. As we can see in Figure 37, one parent chromosome allocates VR instances onto PP1, PP2 and PP3 whilst the other assigns VRs onto PP4, PP5 and PP6. These two parent chromosomes consolidate 6 VRs onto 3 PPs. However, when the crossover happens using the single-point crossover, 4 PPs are needed to host VRs in two children chromosomes. If the migration cost is not considered as the network topology is unknown, from the perspective of power consumption, the two children chromosomes have worse performance than that of parent chromosomes. Thus, the crossover mechanism does not improve the quality of solutions in this example.

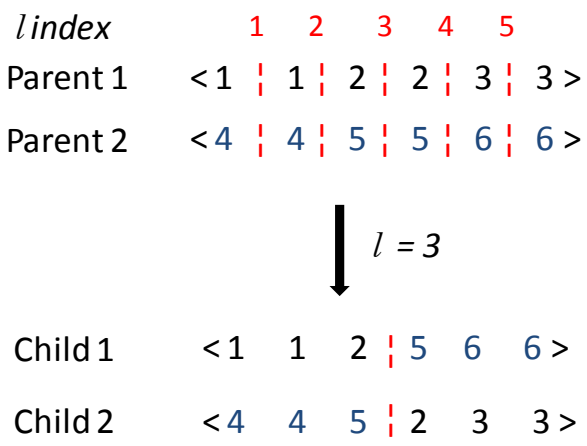

Figure 37. Example Crossover Function

In summary, when considering the computational time and quality of solutions, a low mutation rate is preferred, e.g. 0.1. The crossover mechanisms, i.e. a single-point crossover and BLX-a crossover, have little impact on the performance of VRM_MOEA. Future work could explore either searching for an appropriate crossover mechanism or determining the reasons why the crossover mechanism has limited impact on the destination physical platform selection problem.

5.6.2.2 Maximum Generation

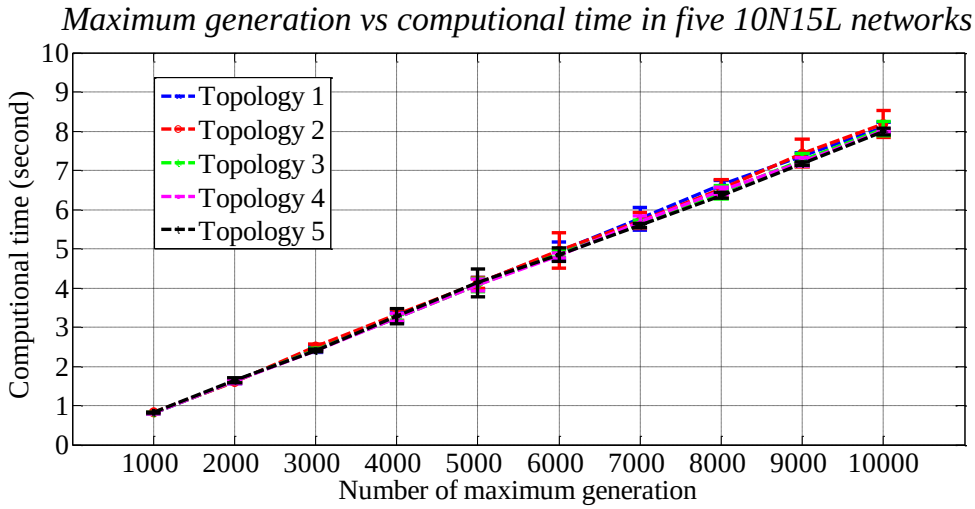

Figure 38. Impact of Topology on Computational Time of 10N15L Networks

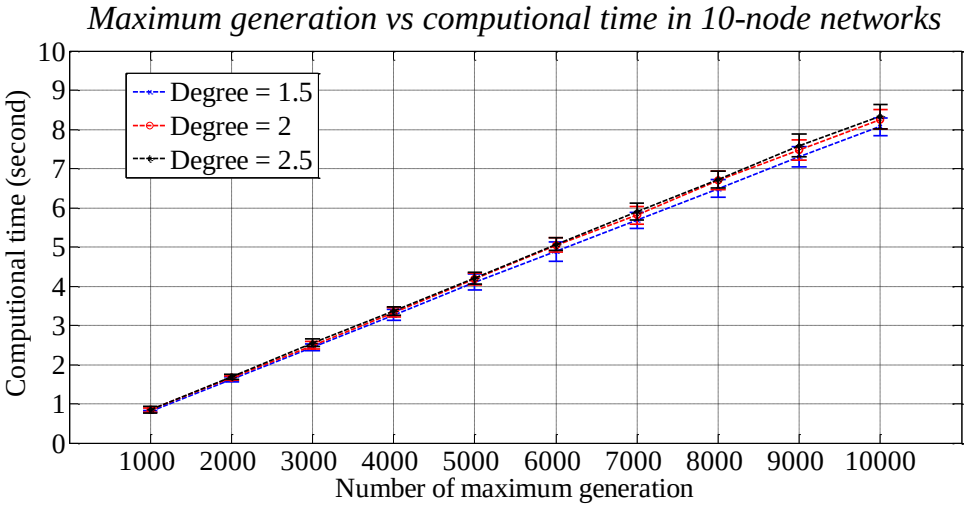

Figure 39. Impact of Average Node Degree on Computational Time of 10-node Networks

In this section, the impact of the maximum number of generations in VRM_MOEA is investigated based on the computational time and the quality of solutions.

Firstly, we explore the computational time with 5 topologies in 10-node networks in Figure 38. For each topology, 30 trials are carried out each with a randomly selected seed. The error bars show the 95% confidence interval of the average computational time.

It is clear that the computational time is similar across the 5 networks with different topologies if the maximum generation value is the same. It means that the network

topology has little impact on the computational time. The computational time also increases linearly with the maximum number of generations.

Secondary, we explore the impact of various network degrees (i.e. average node degrees), i.e. 1.5, 2 and 2.5, on the computational time in Figure 39. For networks with a particular average node degree, three different topologies are used. For each topology, 10 simulation runs with different random number seeds are used. The error bars presents the 95% confidence interval of the average computational time. We can see for a given number of maximum generations, the computational time increases with the average node degree. This is because with a higher network degree, it is possible that more candidate solutions with "good" fitness exist. However, as the size of the secondary population is fixed, not all candidate solutions with "good" fitness can survive. The truncation method in the environmental selection is used to retain a certain number of solutions in the population (as explained in Section 4.4.4, Step 4). Thus, a longer time is needed to do the truncation when more solutions with "good" fitness exist in the secondary population.

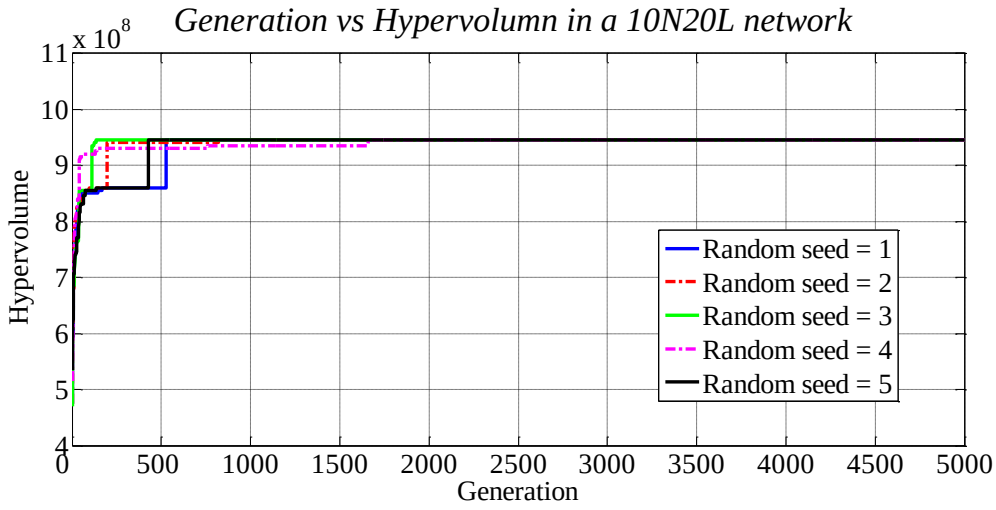

Figure 40. Impact of Random Number Seeds on *Hypervolume* of 10N20L Networks

We also investigate the impact of the number of generations on the value of *Hypervolume*. Firstly, five random number seeds are used in a 10N20L network in Figure 40 and the maximum number of generations is set to 5000. We can see the values of

Hypervolume increase with increasing generations and all of them reach the same value at around 1700 generations, which means the quality of the populations is the same. We call it a "temporary" stable value, as we cannot guarantee that *Hypervolume* has reached its maximum possible value. The "temporary" stable value is the maximum value within 5000 generations.

Then, the results with five 10N20L networks with different topologies are shown in Figure 41. For each topology, 30 separate trials with different random number seeds are used. The trend in *Hypervolume* is similar and the "temporary" stable value is around 2000 iterations. However, for different topologies, the "temporary" stable value is different, though the values are close. This means that with the same average node degree, the network topology has a limited impact on the quality of solutions.

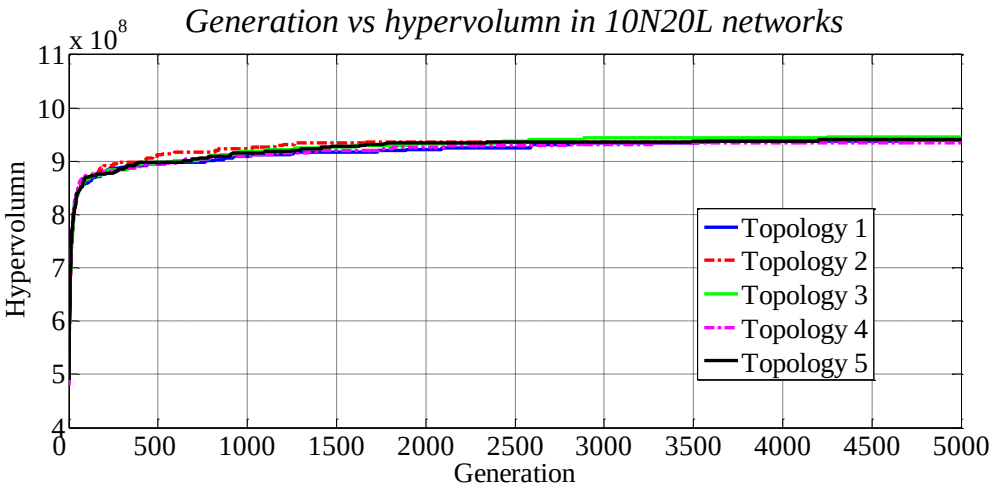

Figure 41. Impact of Network Topology on *Hypervolume* of 10N20L Networks

5.6.2.3 Number of Offspring

The number of offspring indicates the number of qualified individuals generated per generation. In this section, the impact of the number of offspring on VRM_MOEA is investigated. Figure 42 and Figure 43 show the impact of network topology and network

degree in 10-node networks, respectively. The number of trials is 30, each using a randomly generated seed.

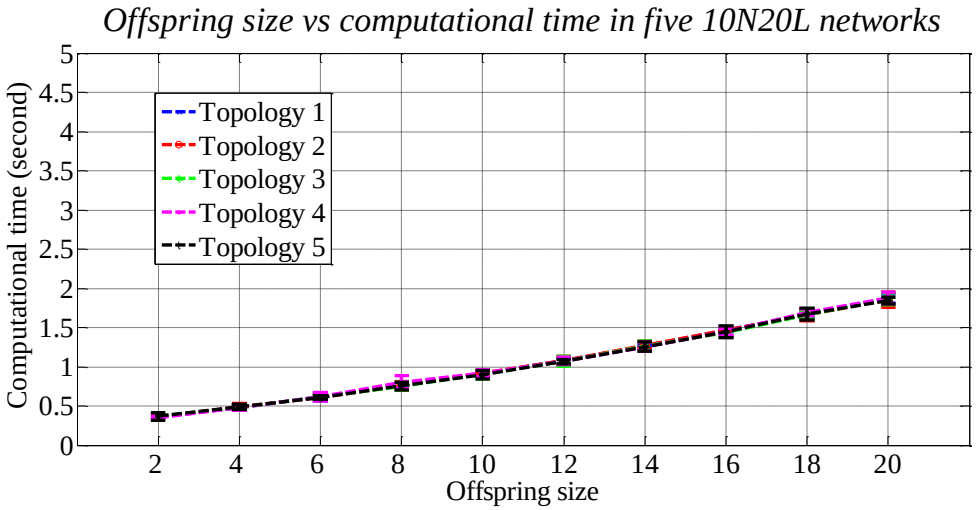

Figure 42. Impact of Network Topology on Computational Time of 10N20L Networks

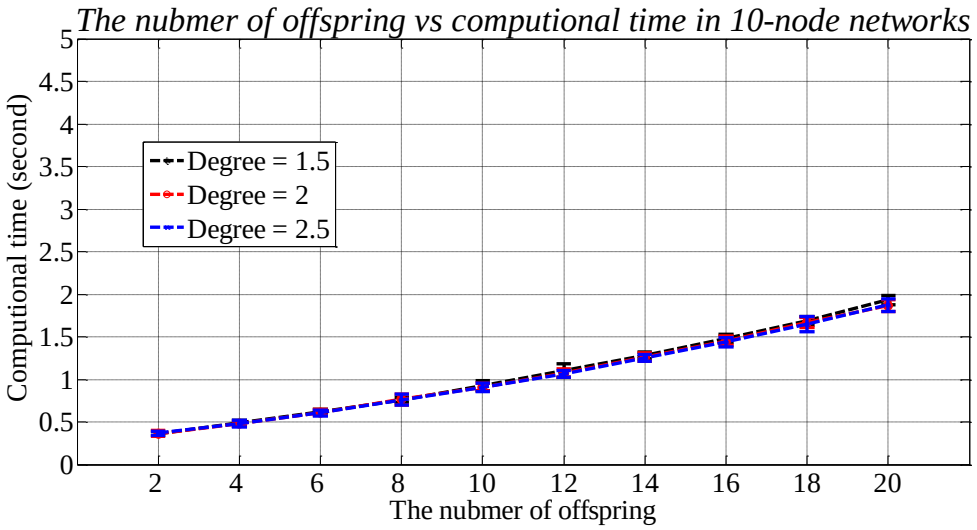

Figure 43. Impact of Network Degree on Computational Time of 10-node Networks

The results show that the network topology and degree have little impact on the computational time. Meanwhile, the computational time increases linearly with the number of offspring. Compared with the computational time of the maximum generation in Figure 38 and Figure 39, the increasing rate associated with the number of offspring is

smaller. For example, when the maximum generation is equal to 10000, the computational time is approximately ten times larger than that of the case when the number of generations is 1000. In the number of offspring case, when the size is 20, the time is around 4 times larger than that of the case when the size equals 2.

We then explore the impact of the number of offspring on the quality of solutions in three 10-node networks with different average node degree; the results are shown in Figure 44. As these three networks may have different groups of good solutions due to the network topology and average node degree, it is not possible to compare their values of *Hypervolume* directly. Hence, we need a new method to compare the impact of different the number of offspring.

As we can see from Figure 40 and Figure 41, when the generation count reaches a certain number, the value of *Hypervolume* becomes "meta-stable". The generation count is denoted as Gen. Although such a meta-stable value cannot guarantee that the value is the optimal one, we can use it as a benchmark. The meta-stable value of *Hypervolume* is donated as H_s. Then, we can explore how many generations it takes for the value of *Hypervolume* to reach H_s with different numbers of offspring. In addition, in order to avoid infinite computation if the value of *Hypervolume* never reaches H_s in a trial, a generation counter, Gen_t, is used as a terminal condition. If the value of *Hypervolume* cannot reach H_s within Gen_t, the simulation stops and Gen is recorded as Gen_t.

Thus, the method is described as follows:

1. Set the maximum number of generations, Gen_m, to be a very large number in order to obtain a meta-stable value as a benchmark. For example, Gen_m is 100000 in the 10-node networks case.
2. Run a simulation to Gen_m and record H_s.
3. Set the value of Gen_t. Note that $\mathrm{Gen_t}$ is smaller than Gen_m. In the 10-node networks case shown in Figure 44, Gen_t is 5000.
4. Run the simulations with different number of offspring. As there are two terminal conditions, Gen is recorded depending on the terminal conditions as follows.

Firstly, when the generation count is smaller than Gen_t and the value of *Hypervolume* is no smaller than H_s, Gen is recorded as the current generation count. Secondary, when the generation reaches Gen_t butthe value of *Hypervolume* is smaller than H_s, record Gen_t as Gen.

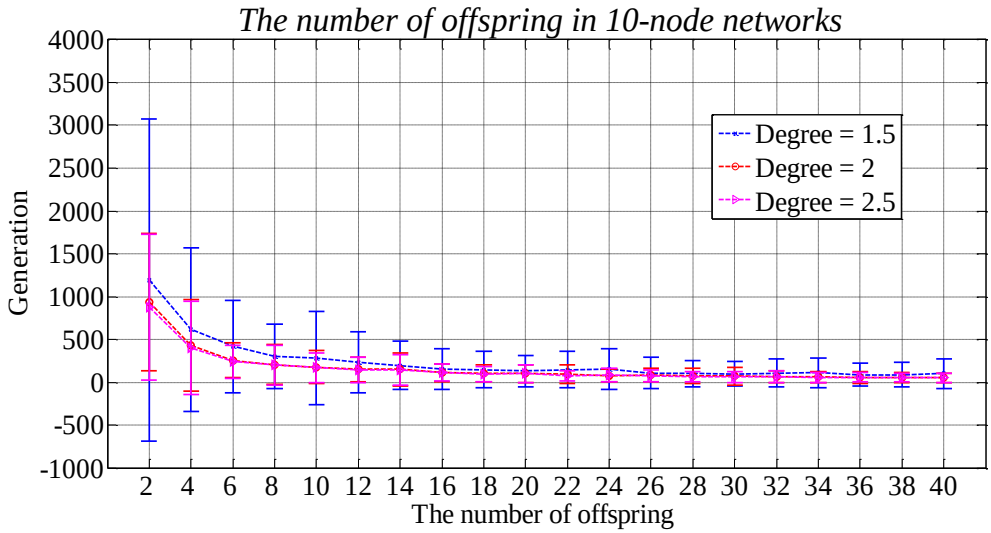

Figure 44. Impact of the Number of Offspring of 10-node Networks

Figure 44 shows that Gen_s decreases with the increasing number of offspring and reaches a stable state when the number of offspring is larger than 22. This means a larger number of offspring can improve the quality of solutions in a given number of generations. From the point of view of network degree, on average, a network with lower network degree needs more generations to reach Gen_s with the same number of offspring. A network with degree 2 and the network with degree 2.5 have similar performance. Compared with these two cases, the network with degree 1.5 requires more generations. This is because a network with lower degree may have fewer solutions with good fitness than that of a network with a higher degree. Thus, the search for such solutions in a network with lower degree is more difficult so more generations are needed.

5.6.2.4 Mating Pool Size

Figure 45 and Figure 46 show that the mating pool size has little impact on the computational time with different topologies and network degree. The mating pool size is

related to the times required to select the candidate solution from the secondary population using tournament selection with replacement (as explained in Section 4.4.4, Step 5). Compared with the viability test and truncation method, the tournament selection mechanism is much simpler. Thus the computational time of the mating selection has little influence on the overall computational time of VRM_MOEA.

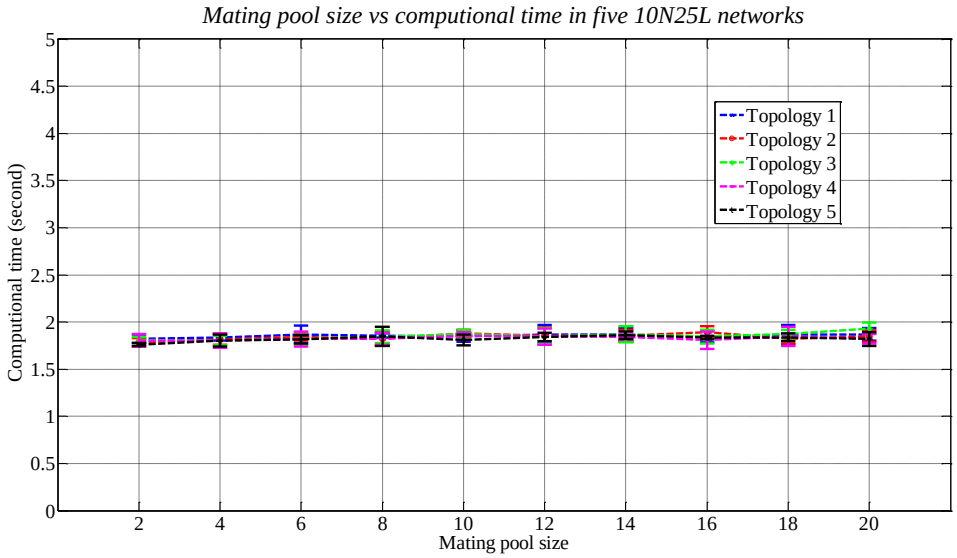

Figure 45. Impact of Mating Pool Size of 10N25L Networks

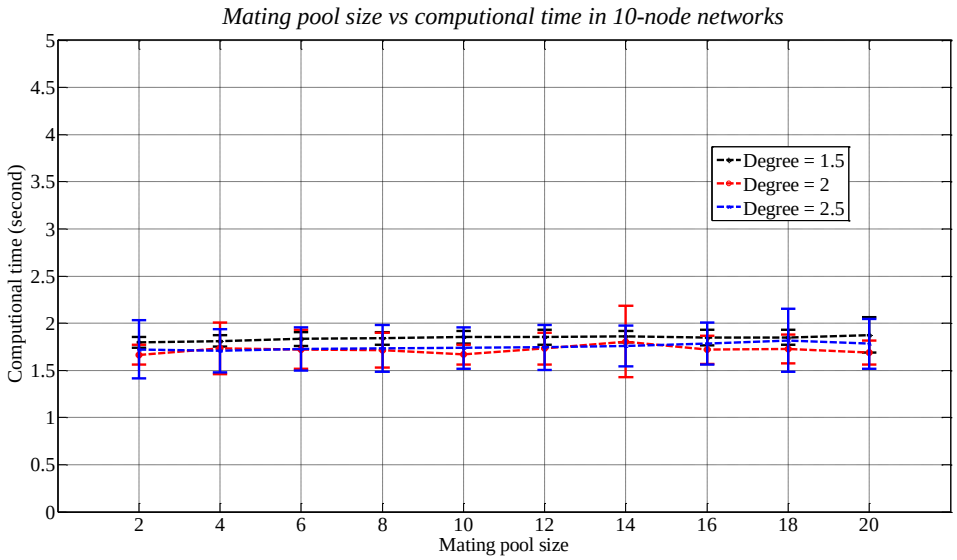

Figure 46. Impact of Mating Pool Size on Computational Time of 10-node Networks

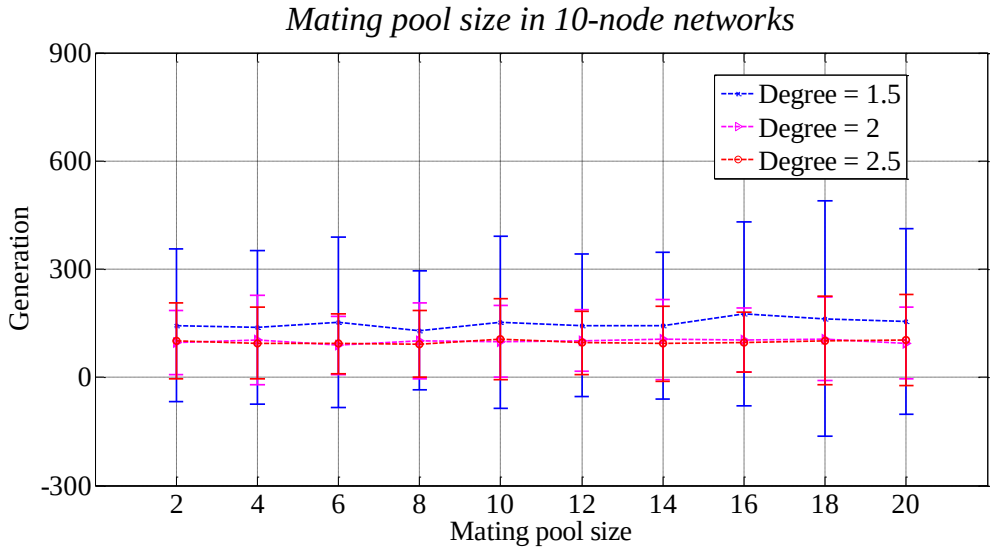

Figure 47. Impact of Mating Pool Size of 10-node Networks

In Figure 47, we used the same method as described in Section 6.2.2.4. The least required number of generations when the value of *Hypervolume* reaches H_s are compared. The results also show that the mating pool size has little impact on the quality of solutions. As we mentioned in Section 6.2.2.1, the one point crossover mechanism has little impact on the algorithm performance and the candidate solutions of the mating pool are mainly used as parent chromosomes in the crossover (remember that the mutation rate should be low). Hence, the crossover mechanism may affect the impact of the mating pool size on the quality of solutions. In addition, similar to the number of offspring in Section 6.2.2.3, a network with low degree may need more generations to approach Gen_s.

5.6.2.5 Traffic Load

In this section, the impact of traffic load on the computational time in 30-node networks with various topologies is explored in Figure 48. It is clear to see that the impact of the topology on computational time increases with the traffic load. Meanwhile, the computational time increases with the traffic load. This is because a busier network has less feasible candidate solutions, as the migration constraints are more difficult to be satisfied compared with that for a network with a light traffic load. Hence, a longer time is needed to search for good solutions.

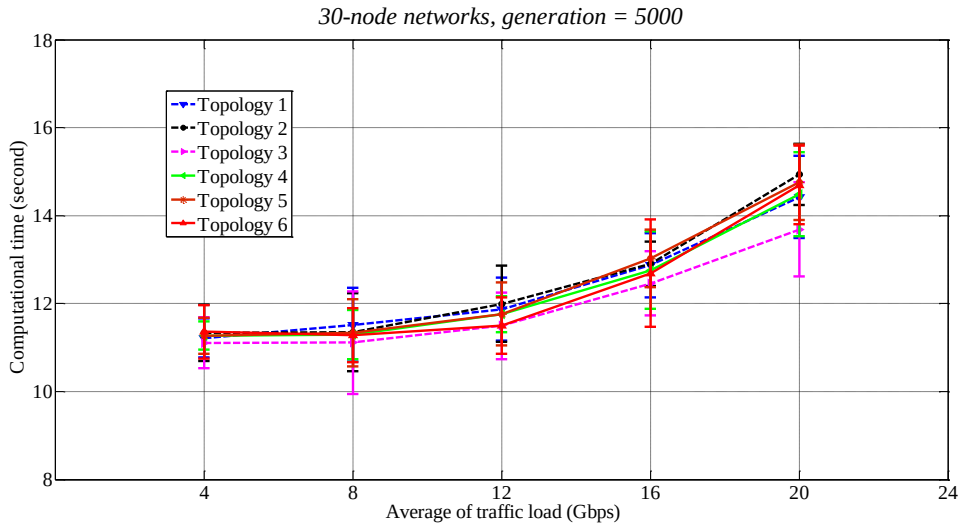

Figure 48. Impact of Traffic Load on Computational Time of 30-node Networks

5.6.2.6 Number of Nodes

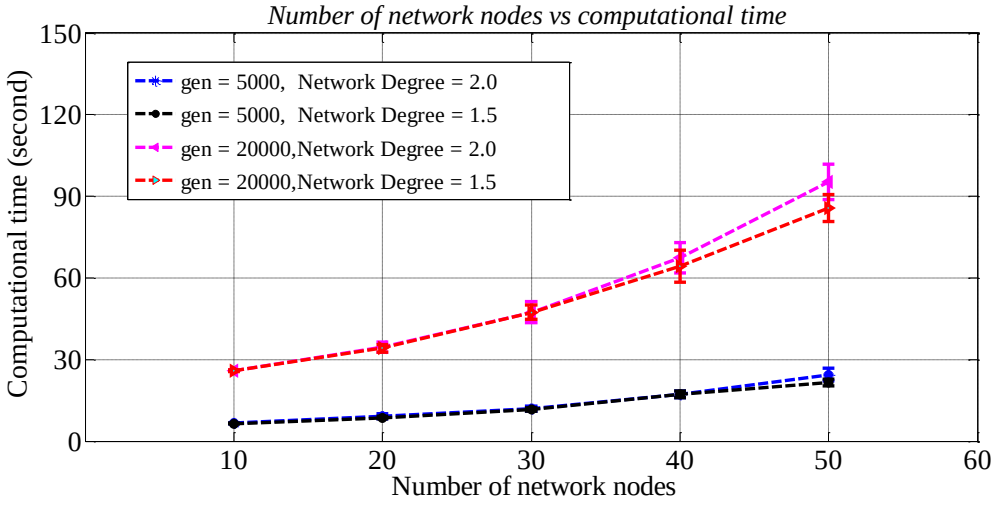

Figure 49. Impact of Number of Nodes on Computational Time

Figure 49 shows the impact of the number of network nodes on the computational time. As shown in Figure 48, the traffic load also influences the computational time when networks with busy traffic (the peak traffic is around 50% - 60% of the PP capacity) are used. For a certain number of network nodes, three network topologies is used in networks with a specific degree and 5 simulations are run with different random seeds.

As expected, the computational time increases linearly with the number of network nodes when the maximum number of generations is set to 5000 and 20000. For a given number of network nodes, the computational time also increases linearly with the generation count. Meanwhile, networks with a higher network degree need more time to perform truncation, which is similar to the results given in Figure 39.

As VRM_MOEA is an online algorithm a group of reasonable solutions need to be obtained in a short period of time, e.g. 5 minutes. It is therefore important to estimate the computational time. Figure 49 provides a means to estimate the computational time of VRM_MOEA if some information is known, e.g. the maximum generation count and the number of nodes. For example, if the number of network nodes is 40 and the maximum generation count is 20000, the computational time will be around 65 seconds. When the maximum generation count is 100000, the time can be estimated as $65 \cdot 5 = 325$ seconds.

5.6.3 **Section Summary**

In Section 6.2, a series of experiments are conducted to investigate the impact of different parameter values on the computational time and the quality of the solutions in VRM_MOEA. Different parameters have different influences on the algorithm. For example, a small mutation rate is appropriate whilst the crossover rate and the mating pool size have little impact on the algorithm performance. In addition, we can see that a high maximum generation count and the number of offspring can help to obtain a group of solutions of good quality. However, based on the simulation results, we cannot summarize a generic method to configure appropriate values of these two parameters across any network. Future work would be needed to determine more precisely the relationship between these two parameters and the performance of VRM_MOEA. As the computational time is a limiting factor, one solution is to select a high maximum generation count and the number of offspring as long as the computational time is within the allowed period.

5.7 **Validation of Simulation**

A simulation model is an approximate imitation of a real system, so validation of the simulation model is required. A simple 3-node-3-link network example is used to validate the entire dynamic energy management framework and details are given in Appendix B, Part 5.

5.8 **Summary**

In this chapter, a simulation model is developed to evaluate the performance of the dynamic energy management framework. The simulation is a stochastic and discrete event simulation using C/C++ with Microsoft Visual Studio 2012. The data traffic is modelled as time-varying flows and signalling information is modelled as discrete packets.

Initially, the details of the simulation tool are described, e.g. entities, events and components. Then, several important functional modules are described such as the network topology, the traffic model and the random number generator. In the network topology section, three network topologies are presented as well as a simple network topology generator. The simple random network topology generator is developed for exploring the relationship between the performance of framework and the average node degree. Two types of traffic model are used in our study, i.e. (1) a sinusoid-like function with some random noise and (2) an historical traffic dataset of the Abilene network. The random number generator is then introduced.

Next, we investigated the appropriate parameter settings within VRM_MOEA. Based on simulation results, the mutation rate is recommended to be low, e.g. 0.1. Meanwhile, the size of mating pool and the crossover rate have little impact on VRM_MOEA performance. This may be due to the nature of destination physical platform selection problem or an inappropriate crossover mechanism. More investigations are needed in the

future. Furthermore, the number of offspring and the maximum generation count should be large as they can help to obtain a group of good quality solutions. However, based on the simulation results, we cannot generalise and provide a method to select appropriate values of these two parameters for any particular network. More effort is needed to examine the relationship between these two parameters and the performance of VRM_MOEA. Finally, the validation of simulation is described.

CHAPTER 6 SIMULATION RESULTS AND DISCUSSION

6.1 Introduction

Chapter 5 describes the simulation modelling of the dynamic energy management framework and parameter setting in VRM_MOEA. The next step is to measure the performance of the framework using the simulation tool. Hence, Chapter 6 presents an investigation of the performance of framework including simulation results with associated discussions. The remainder of Chapter 6 is organized as follows. Initially, some simulation parameters and assumptions are described in Section 6.2; then simulation results with various parameters are shown in Section 6.3, e.g. the weighted migration cost component parameters and the *Quiet* and *Busy* threshold settings. Finally, a chapter summary is presented in Section 6.4.

6.2 Simulation Parameters and Assumptions

The computing environment used for the simulations was a Microsoft Windows 7 Enterprise, Intel® Core™ i5 CPU 661 @ 3.33GHz, with 4G of RAM. Note that in the simulations, "6N8L", "14N21L" and "Abilene network" indicate the three reference network topologies described in Section 5.3.1. The remaining network topologies, e.g. networks with average network node degree of 1.5 or 10N25L, are generated by the simple Random Network Topology Generator (RNTG, as described in Section 5.3.2).

Some simulation parameters and assumptions are described as follows.

1. Given an optical network of a graph $G = (V, E)$ where V is the number of vertices and E is the number of edges in the network. Optical Line Amplifiers (OLAs) are placed on the network edges and regenerators are placed on the network node. In our framework, each node is composed of a PP and an ROADM. Thus, for a $V - node - E - Link$ network, it contains V PPs, V ROADMs, V VRs and E links.
2. We assume that all PPs are homogeneous. If PPs have different parameters, such as fabric switch capacity and the number of line-cards, these parameters are considered as restrictions in VRM_MOEA. For the sake of simplicity, all PPs have the same configuration.
3. The capacity of a line-card is 40Gbps and each line-card comprises a single 40Gbps bi-direction interface. The switch fabric capacity of a PP supports the packet switching amongst various line-cards.
4. ROADMs are connected by a pair of transmission fibres. The number of wavelengths per optical fibre is $W = 40$. The transmission rate of each wavelength is 40Gbps.
5. The degree of ROADMs is determined by the network topology. For example, if a ROADM has two neighbors, the degree is 2. This means that the ROADM can add, drop and bypass any wavelengths over these two directions. For each direction, ROADMs can deal with a transmission fibre pair with 40 wavelengths.
6. In VROOM [44], the migration duration of a virtual router instance is about 33.9 seconds with a software data plane entity, which has 15k routes. Hence, we assume that the duration of virtual router migration event is 34 seconds.
7. The maximum length without the need of OLA is $\varepsilon_{OLA} = 80km$.
8. The scenario without VRM capacity is called the Baseline configuration. A scenario with VRM capability is called an Energy-Efficiency (EE) configuration.
9. The network state information is collected every 15 minutes. This is because 15 minutes is long enough to calculate VRM_MOEA and move the virtual router

instance to their destination physical platform, but short enough to capture the essential mid-term traffic dynamics [95].

10. The parameters associated with the power consumption model are shown in Table 4.

Table 4. The Parameters Associated with Power Consumption Model

Name	Typical value	Description	Ref.
M_{lc}	16	The maximum number of line cards each LCC supports	[163]
M_{lcc}	9	The maximum number of LCCs each FCS supports	[163]
P_{lc}	500W	The power consumption of a line-card	[165][166]
P_{lcc}	1630	The power consumption of a LCC with 16 line card slot	[167]
P_{fcc}	7036W	The power consumption of a FCC	[168]
θ	0.05	A small fraction of the active base system power consumption when a PP is in the sleep state	Self-defined
$P_{ROADM}^{40}(\omega)$	$\omega \cdot 135 + 150$	The power consumption of an ROADM with 40 channels. ω is the node degree.	[109]
e_{OA}	110W	The power consumption of a long span optical line amplifier	[109]

11. The parameters associated with VRM_MOEA are shown in Table 5 for three reference networks.

Table 5. The Parameters Associated with VRM_MOEA

Parameter	**6N8L**	**14N21L**	**Abilene network**
Length of a solution	6	14	10
Primary population size	20	30	25
Secondary population size	10	10	10
The number of offspring	20	30	30

Mating pool size	5	10	10
Crossover rate	0.1	0.1	0.1
Mutation rate	0.1	0.1	0.1
Maximum generation	2000	20000	10000
w_{cost}	0.5	0.5	0.5
Simulation length	5 days	5 days	7 days
Random number seeds	30	10	10
Busy Threshold	0.8	0.8	0.8
Quiet Threshold	0.3	0.3	0.3

6.3 **Numerical Simulations**

In this section, we explore the impact of various parameters, e.g. traffic load and thresholds on the performance of the framework.

6.3.1 **Impact of Traffic Load**

A set of experiments with varying traffic load between each node pair in the 6N8L network and the 14N21L network have been carried out to evaluate the impact of the traffic load on the dynamic energy management framework. The results are shown in Figure 50 and Figure 51, respectively. The simulation length is 5 days. A one-day "warm up" period is used to allow the simulation to reach a stable state, so the first day data is excluded from the results. As the results are similar from Day 2 to Day 5, the energy savings per hour during Day 5 are shown.

Note that the traffic load in Figure 50 and Figure 51 is the maximum traffic demand in the fluid flow traffic model (as explained in Section 5.4). As we use PPs with same configuration in the two networks and the PPs need to be able to process the traffic demand in the Baseline configuration, the maximum traffic demand that the two

networks can sustain differ. Hence, the 6N8L network in Figure 50 can sustain a higher traffic demand than that of 14N21L network in Figure 51.

Based on the traffic model (as explained in Section 5.4), the peak hours are around hour 14 to hour 20 and off-peak hours are hour 02 to hour 08. Then Baseline configuration provides an upper energy consumption limit per hour. Hence, the percentage energy saving per hour is calculated as follows:

$$Energy\ saving\ \% = \frac{Energy_{Baseline} - Energy_{EE}}{E_{Baseline}} \times 100 \quad (6.1)$$

where $Energy_{Baseline}$ is the energy consumption of Baseline configuration within an hour and $Energy_{EE}$ is that of an EE configuration. For the sake of simplicity, the average energy saving value is shown in Figure 50 and Figure 51.

In Figure 50 and Figure 51, it is clear to see the energy saving characteristics in the 6N8L and 14N21L networks with various traffic loads follow a similar trend; the energy saving reaches a maximum value during off-peak hours and a minimum value during peak hours as the energy consumption per hour fluctuates with changes in the traffic load.

Figure 52 and Figure 53 show the total energy saving with different traffic loads. As expected, the percentage energy saving decreases with increasing traffic load. It implies that in a busier network, it is more difficult to obtain energy savings because there are fewer opportunities to consolidate VR instances onto fewer PPs.

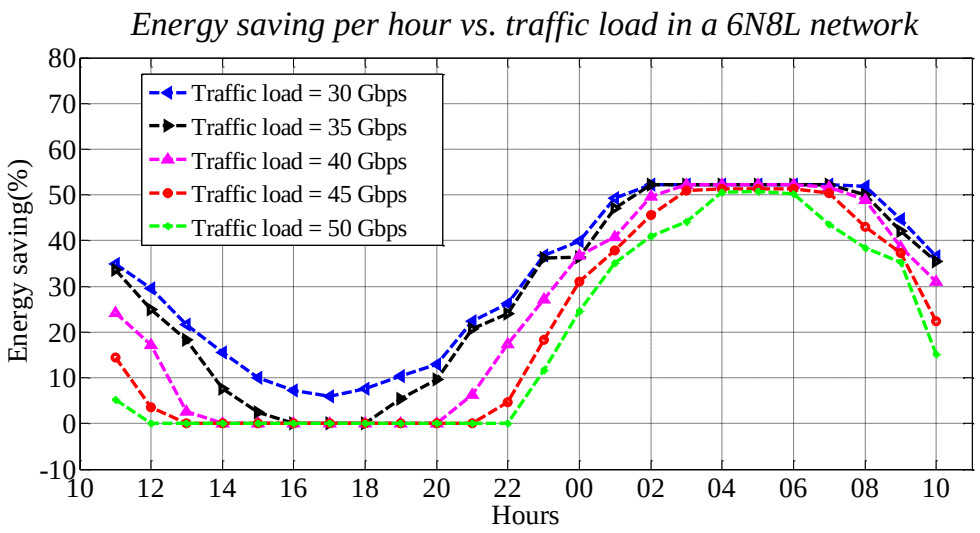

Figure 50. Energy Saving per Hour with Different Traffic Load in the 6N8L Network

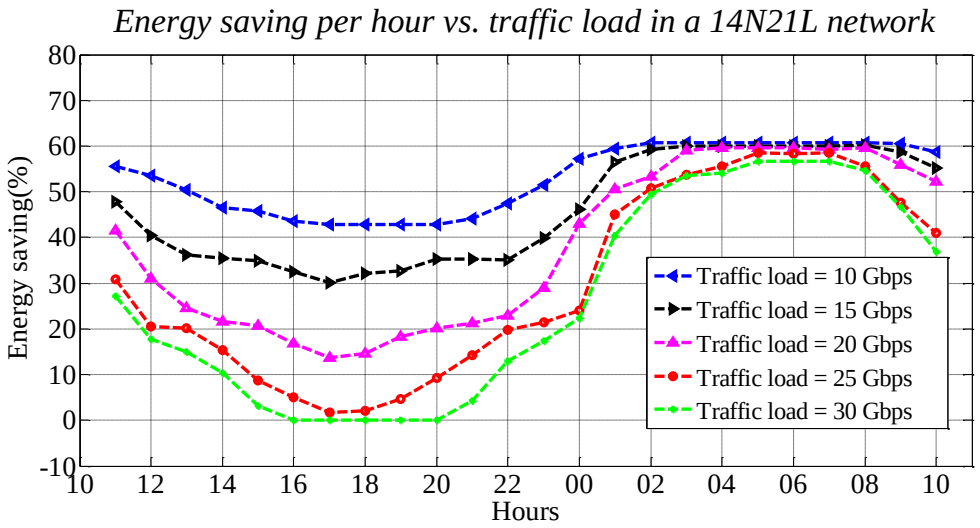

Figure 51. Energy Saving per Hour with Different Traffic Load in the 14N21L Network

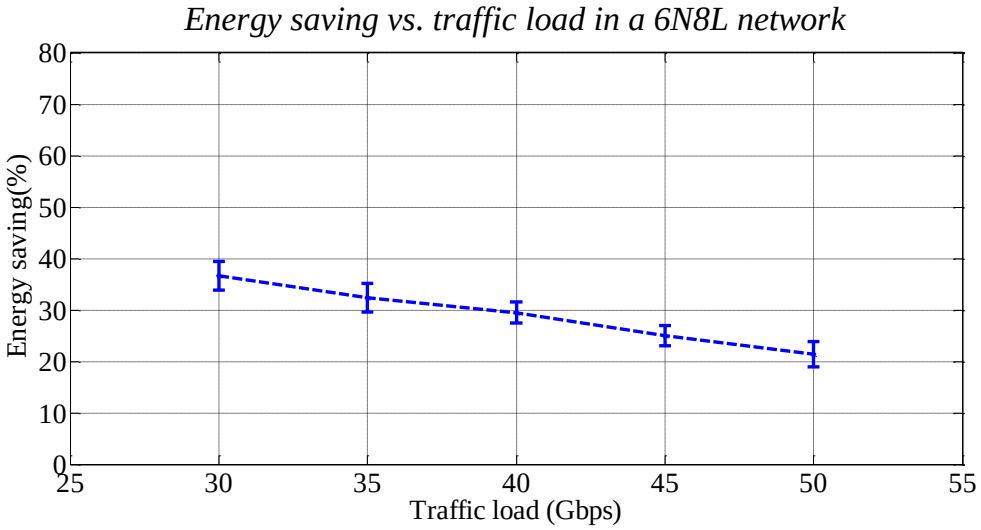

Figure 52. Total Energy Saving with Different Traffic Load in the 6N8L Network

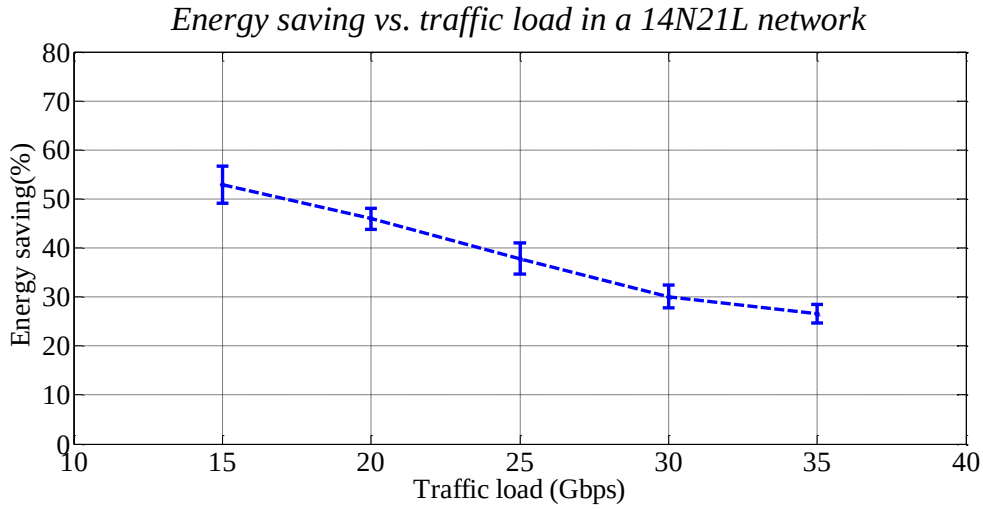

Figure 53. Total Energy Saving with Different Traffic Load in the 14N21L Network

A set of experiments is performed in the Abilene network with historical traffic data of the 1st week, 13th week and 24th week. The traffic loads during these weeks are shown in Appendix B, Part 2. As the date information is anonymous, we can only guess the five weekdays and two weekend days based on the traffic load. For example, we can see that the traffic load of the 1st week has a clear daily pattern. On average, the traffic load from Day 1 to Day 5 is larger than that of Day 6 and Day 7. Hence, we assume the first five

days are weekdays and Day 6 and Day 7 are weekends. From the view of energy saving, the energy saving varies with the traffic variation in the first five days. As the network is quieter in the Day 6 and Day 7, which are weekends, no VRM event happens and energy consumption remains the same.

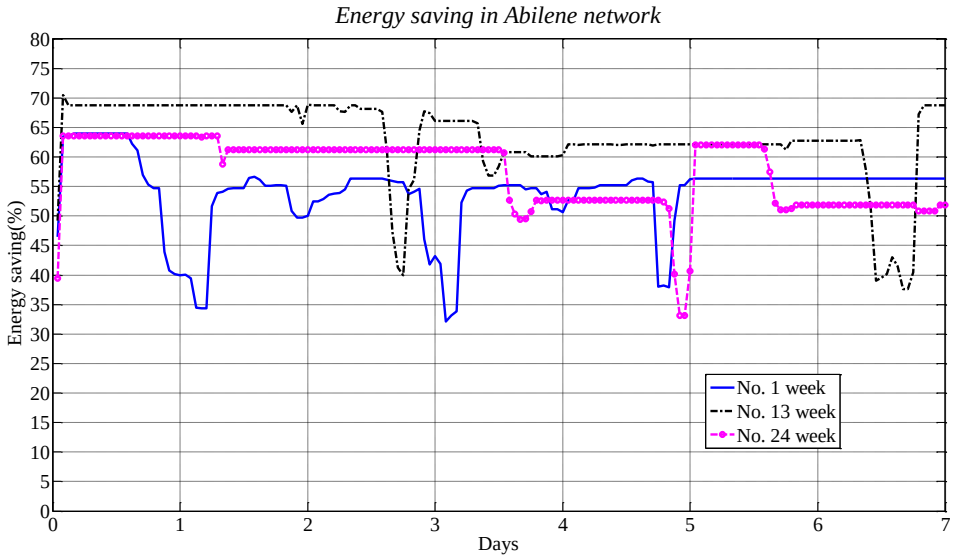

Figure 54. Total Energy Saving per Hour of Three Weeks in the Abilene Network

It should be noted that the traffic load has some spikes, which are double or triple that of the peak traffic load the rest of the time. In order to handle these traffic spikes, we increased the capacity of PPs until they can process these traffic surges in Baseline configurations. However, the rest of the time such configurations lead the PP workload to be low and the PPs are effectively in their *Quiet* state. Hence, VR instances are consolidated onto fewer PPs for saving energy in EE configurations. When a VRM event completes, these working PPs subsequently remain in their normal state. If there is no big traffic variation, these VR instances keep running on the same PPs. This explains why the energy saving per hour shows a flat line in the 13th week and 24th week though they also exhibit a daily pattern to some extent.

Figure 54 also shows that if the traffic load changes significantly in a short time (e.g. spikes), the framework cannot respond quickly. As a reactive mechanism is used in the

framework, when the CCU collects the updated network state information, it still needs a period of time to compute VRM_MOEA and allow for the VR instances to be moved to their appropriate destinations. During such a time, some PPs cannot sustain the traffic and traffic loss may happen. Moreover, such a significant change may be unpredictable, so a proactive mechanism may not help in such a scenario. One possible solution is is to use a combination of buffering and possibly lowering the *Busy* threshold.

6.3.2 **Impact of Busy and Quiet Thresholds**

We investigate the effect of adjusting the *Quiet* and *Busy* thresholds in the 6N8L network. The results are given in Figure 55. The range of the *Quiet* threshold is [0.1, 0.7] and the step interval is 0.1. The range of *Busy* threshold is [0.3, 0.9] and the step interval is 0.1. The *Busy* threshold cannot be 1 as a safety margin is required. There are two constraints when setting the two thresholds. Firstly, the *Busy* threshold is larger than the *Quiet* threshold. Then, if the difference between two thresholds is small, a PP may change its state frequently which may cause VR instances move around the network frequently. Hence, we assume that there is at least 20% difference between the *Quiet* and *Busy* thresholds.

Figure 55 shows the energy saving with different *Quiet* and *Busy* thresholds. The different colours represent the different values. The crimson colour stands for a highest value in the figure. Hence, the largest energy saving value is around 30%. When the *Busy* threshold is larger than 0.6 and the *Quiet* threshold is larger than 0.4, the energy saving is largest.

It is clear that with a given *Busy* threshold the energy saving initially goes up linearly with the increasing *Quiet* threshold and then reaches a stable state. For example, when the *Busy* threshold is equal to 0.9, the energy saving increases firstly. When the *Quiet* threshold is larger than 0.4, the energy saving remains constant. This is because at first, a configuration with higher *Quiet* threshold allows PPs to remain in their *Quiet* mode for longer periods which leads to greater energy saving. However, as PPs have limited switch

fabric capacity, they are not able to host more than a certain number of VR instances. When the PP capacity is reached (or the PP cannot host another VR instance), VR instances cannot be consolidated onto fewer PPs even if these PPs are in their Quiet state. Thus, the energy saving remains the same. A similar trend exists when observing the *Busy* threshold with a fixed *Quiet* threshold.

Note that a higher *Busy* threshold brings with it a higher risk of traffic loss. If traffic is bursty and increases quickly before a PP enters the *Busy* state close to the switch fabric capacity, the PP switch may become overwhelmed and incur traffic loss before VRM can take place.

Figure 56 shows the additional optical hop required in the 6N8L network with different *Quiet* and *Busy* thresholds. The average optical hop count of traffic in the Baseline configuration is 2.2778. By applying VRM approach, the logical IP-layer topology remains unchanged whilst the optical layer is adjusted to allow the traffic to be forwarded appropriately to each VR's current location. When the traffic goes through an ROADM, the hop count is incremented by one. In an EE configuration, the traffic is transmitted along a longer optical path if a VR instance is not running on its default PP location. Hence, it is clear to see that EE configurations with different *Quiet* and *Busy* thresholds result in additional optical hops.

In order to select appropriate *Quiet* and *Busy* thresholds settings, both the energy saving and the additional optical hop count need to be considered. However, we can see that bigger energy savings usually require higher optical hop counts. This is because as more VR instances are consolidated onto fewer PPs, the traffic needs to be transmitted on a longer optical path to the remotely located VR.

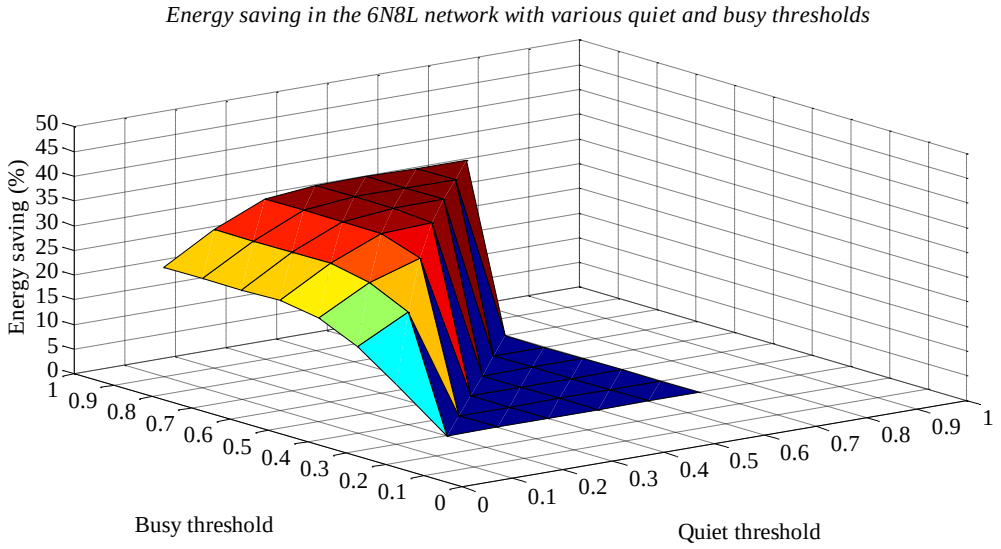

Figure 55. Total Energy Saving of Different Threshold Settings

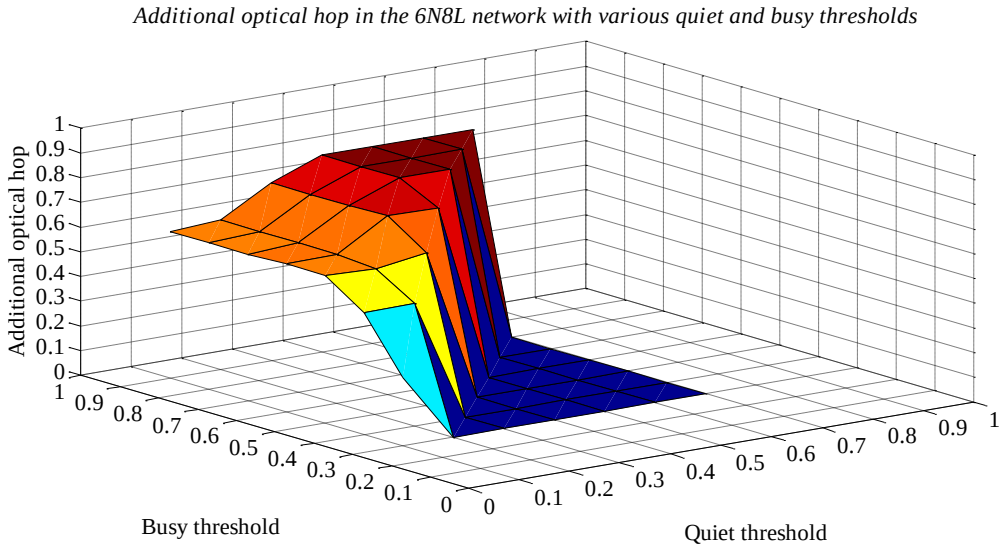

Figure 56. Additional Optical Hop in the 6N8L Network with Different Threshold Settings

6.3.3 Impact of Migration Cost Setting

In Figure 57, we explore the number of occupied optical channels in the Baseline configuration and 5 EE configurations with different parameter weightings in the 14N21L network. The number of occupied optical channels is recorded every 15 minutes

during the fifth day. For the sake of simplicity, only the average number of occupied channels is shown.

The results include a Baseline configuration (Configuration 1) which yields the minimum number of optical channels used as no migration is permitted, and 5 EE configurations with the two-part cost weightings as given in Table 3. The migration cost function is expressed as Cost (w_{cost}, $1 - w_{cost}$), where w_{cost} stand for weighted coefficient of two migration cost components, i.e. *Cost_a* and *Cost_b* (as explained in Section 4.4.4 Step 2).

The first term, *Cost_a*, is based on the hop-distance from default PP locations to destination PP locations; the second is the hop-distance from current PP locations to destination PP locations. In Configurations 2 and 6, only one cost component is considered in VRM_MOEA. In Configuration 2 and 3, w_{cost} is smaller than $(1 - w_{cost})$ indicating that the *Cost_b* term is more important than the *Cost_a* term. The two terms are equivalent in Configuration 4. The *Cost_a* term is more significant than *Cost_b* in Configuration 5 and 6.

Table 6. Configurations

Configuration number.	Configuration Setting
1	Baseline, No VRM
2	Cost (0.0,1.0)
3	Cost (0.25,0.75)
4	Cost (0.5,0.5)
5	Cost (0.75,0.25)
6	Cost (1.0,0.0)

Figure 57 shows that the number of occupied optical channels in Configuration 1 fluctuates with the daily traffic load as additional channels are required to transport traffic between specific ingress and egress points during peak times (i.e. each channel is limited to 40Gps and the traffic load can exceed this value). The number of occupied optical channels for all the remaining configurations is higher than that of Configuration 1. This

is because when VR instances are moved to remote PPs leaving their default PPs, more optical channels are needed for transmitting the traffic to the new locations.

Amongst Configurations 2 to 6, those with a lower *Cost_a* value result in a higher number of occupied channels. The *Cost_a* term represents the hop distance from the default PP to the destination PP. When the weighting of *Cost_a* increases, the algorithm searches for solutions that are close to a VR's default location which reduces the number of additional optical channels from the default PPs to the destination PPs. Configuration 2 is the worst because *Cost_a* term is zero and the hop-distance from a default PP and a destination PP is not considered in VRM_MOEA.

Note that at peak times, the number of occupied optical channels reduces in Figure 57. This is because when the network is busy, VR instances are distributed across more PPs and some of them may go back to their default PP locations. Hence, some optical channels which are used to transmit the traffic to remotely located VR instances in previous configuration are no longer needed. These optical channels are released.

Figure 58 shows the number of occupied optical channels in the Abilene network in the first week with different cost settings (the results of remaining two weeks are shown in Appendix C, Part 2). The results are similar to the 14N21L network. Based on this observation, in order to decrease the number of occupied optical channels, it can be deduced that it is necessary for the *Cost_a* weighting to be higher than that of the Cost_b term.

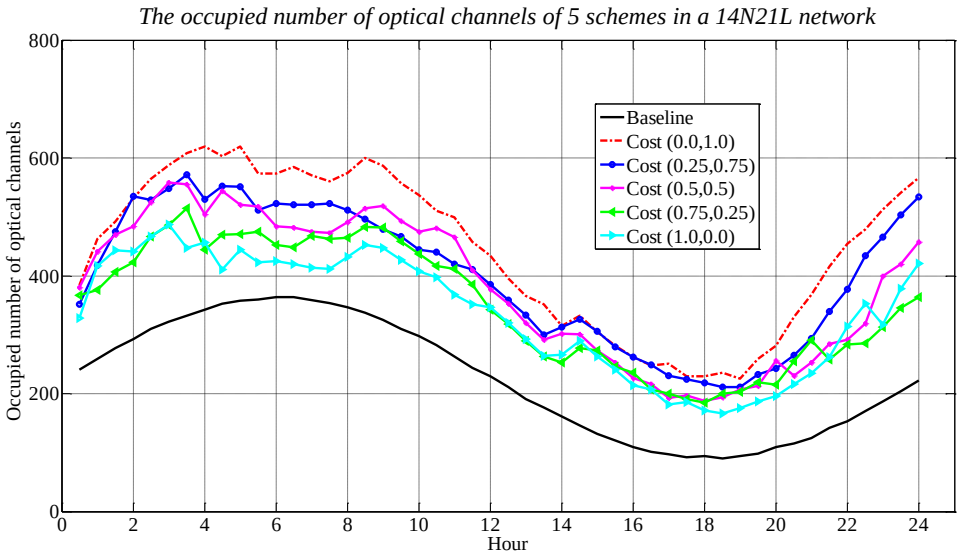

Figure 57. Number of Occupied Optical Channels of Different Configurations in the 14N21L Network

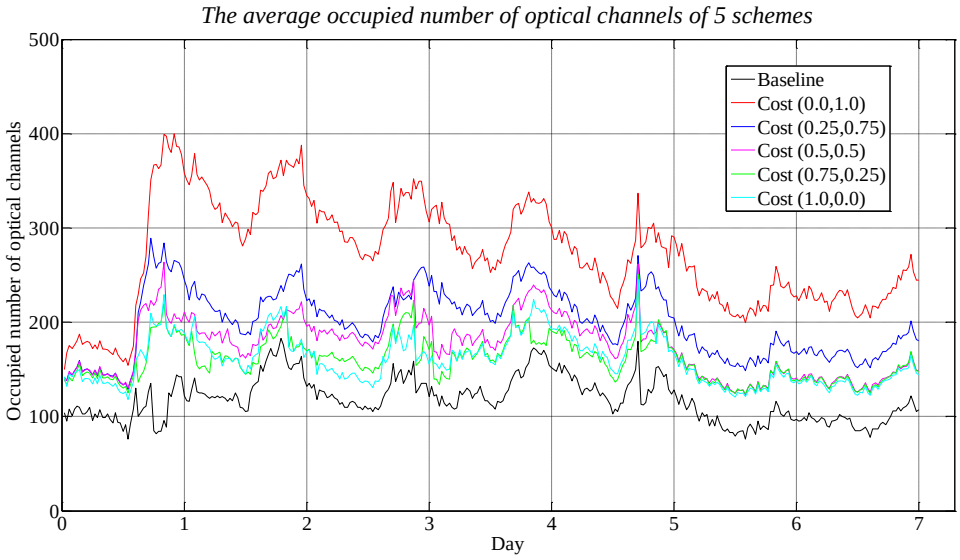

Figure 58. Number of Occupied Optical Channels of Different Configurations in the Abilene Network

6.4 **Summary**

In this chapter, we have explored the impact of different parameters on the performance the dynamic energy management framework. A series of experiments are conducted to investigate the impact of traffic load, *Quiet* and *Busy* thresholds, and the migration cost on the framework.

Firstly, in the simulations with various traffic loads, the percentage of energy saving decreases with the increasing traffic load, as expected. It implies that in a busier network, it is more difficult to obtain the energy saving, as there are fewer opportunities to consolidate VR instances onto fewer PPs.

In the 6N8L network case, when the *Busy* threshold is larger than 0.6 and the *Quiet* threshold is larger than 0.4, the energy saving is largest. However, at the same time, additional optical hops are required as the traffic must be transmitted to the remotely located VR instance(s) via longer paths.

Finally, we investigated the impact of different migration cost parameter weightings. In order to decrease the number of occupied optical channels, it is necessary for the *Cost_a* weighting to be higher than that of the *Cost_b* term.

CHAPTER 7 CONCLUSIONS AND FURTHER WORK

7.1 **Conclusions**

Improving the energy efficiency of the Internet has become an important topic in recent years as energy consumption may become one of the main constraints for the further growth of the Internet.

The energy consumption of the Internet is continuously rising which essentially depends on the increasing amount of traffic it supports. This increasing demand is driven by a significant increase in the customer population and ongoing development of various forms of Internet-based services. Meanwhile, the improvements in silicon technologies are not enough to cope with these network energy efficiency issues. Thus, in order to cut the operational costs of ISPs, reducing the carbon dioxide emissions associated with electricity generation and lowering the power/energy density of network equipment, new energy efficient approaches need to be developed.

Given the above background, an evolutionary algorithm based dynamic energy management framework is proposed for IP over DWDM core networks in this thesis. Although core networks represent a small fraction of the total energy consumption of the Internet, they are gradually ascending the energy consumption "league" due to the increasing popularity of bandwidth-intensive services.

The main concept underpinning our framework is to combine infrastructure sleeping and virtual router migration to enable a network to operate in a more energy efficient manner. Infrastructure sleeping enables unneeded equipment to be switched off or put into a sleep state during off-peak periods. Typically, the traffic demand of core (Tier 2/3)

networks has a regular diurnal pattern based on people's activities. Infrastructure sleeping exploits the characteristics of the traffic and hence has become one of the most promising energy saving mechanisms.

However, infrastructure sleeping suffers from a crucial problem. As routers or links are put to sleep, they lose the ability to exchange routing protocol signalling messages with other routers. Subsequently, the logical IP-layer topology is changed, which triggers a series of reconvergence events. Such events can cause network discontinuities and disruption. Hence, a key constraint of the framework is to maintain the logical IP-layer topology whilst routers are in the sleep state during off-peak hours. In order to avoid the topology change problem, Virtual Router Migration (VRM) is used in the framework. VRM allows a VR instance to freely move between Physical Platforms (PPs) avoiding unnecessary changes to the logical topology. As no logical topology changes happen when PPs enter or leave their sleep state, no network discontinuities or service disruption arises.

In order to realise the new energy management framework, some additional requirements of the network architecture and associated functionality need to be satisfied. For example, a centralized control unit is needed to collect the network state information and a PP has to possess a sleep capability

After two important approaches, i.e. infrastructure sleeping and VRM, and some requirements are described, two significant issues are considered given various constraints, namely: when to trigger VRM and where to move to VR instances to?

For the first issue, a reactive mechanism is used to trigger VRM by monitoring the network state. Typically, the network state is monitored at 15-minute intervals. If the network state satisfies the migration condition, the destination physical platform selection algorithm is invoked for choosing the appropriate locations for VR instances. The destination physical platform selection algorithm is an evolutionary-based algorithm called VRM_MOEA. VRM_MOEA considers different constraints, e.g. the PP capacity and the migration cost. A new individual representation, two migration constraints and

dual objective functions are developed within the algorithm. VRM_MOEA needs to consolidate functionality of suitable VR instances onto fewer PPs during the off-peak hours and distribute co-existing VRs away from busy PPs during peak hours.

Moreover, as VR instances are consolidated onto fewer PPs during off-peak hours for energy saving, some additional optical connections are needed for forwarding the traffic to the remote VRs responsible for processing the packets. These new optical connections also ensure logical separation of the different VR instances that co-exist on the same PP. Hence, a new automatic optical connection management scheme is also described in the framework.

The overall dynamic energy management procedure is described as follows. The first step is to collect and analyze the network conditions. If the network state matches the migration conditions then VRM_MOEA is invoked to select the appropriate destination PPs for the VRs. Then, after a group of reasonable solutions are obtained in VRM_MOEA, the optical resource availability is tested. If the available optical resources can support the identified solution, new optical connections are established for VRM and the corresponding PP(s) are switched on in preparation for hosting the new VR instances. Next, the VR(s) are moved to their appropriate destination(s) based on the identified solution. Where appropriate, the corresponding PP(s) (i.e. no longer hosting VRs) are switched off and the unneeded optical connection(s) are released. Then, the cycle returns to the first step to check the network state.

Once the dynamic energy framework is described, the next step is to measure the performance of the framework. Because of this, a new hybrid simulation platform is developed which is able to capture the functionality of the optical layer, the IP layer data-path and the IP/Optical control plane. The optical layer represents optical channels via a discrete set of integers with associated state information. Meanwhile, a fluid-flow model represents the IP data-path and the IP/optical control plane models the signalling messages using discrete packets. The reason a hybrid environment is required is that it is

not viable to model IP data-flows as discrete streams of packets on an optical infrastructure over many hours.

We then explore the performance of VRM_MOEA as well as the dynamic energy management framework for different parameter values. We tried to ascertain the appropriate parameter values in VRM_MOEA through experimental comparison. The VRM_MOEA results show that a low mutation rate is preferable, e.g. 0.1. Meanwhile, the size of the mating pool and the crossover rate has little impact on the performance of VRM_MOEA in medium-size networks (e.g. 10-node networks). In addition, we can see high values for the maximum generation count and the number of offspring can improve the quality of solutions and then when they reach a certain value, the quality of solutions stabilizes. However, further work is needed to predict the appropriate value of these two parameters for any given network.

Simulation results of the whole framework show that the ability of the dynamic energy management scheme to save energy depends on many factors, such as *Quiet* and *Busy* thresholds and traffic load. In addition, more optical channel resources are needed in order to transmit the traffic to remotely located VRs. By adjusting the migration cost terms we can influence the selected migration location. We can influence the solution in order to keep migrations close to the default physical platform location and/or close to the site currently accommodating the virtual router(s).

7.2 **Further Work**

Several directions are worthy of further research, as described as follows:

1. In the Section 5.6, we examine whether appropriate parameter settings can be predicted for VRM_MOEA for various network scenarios, but only limited success is achieved. There remain questions that need to be explored. For example, why do the one point crossover and Blend crossover mechanisms have little impact on the overall performance of VRM_MOEA. This could be due to the

nature of the destination physical platform selection problem or these two crossover mechanisms are particularly not suitable for VRM_MOEA. Another question is how to devise a suitable method to obtain appropriate values for the maximum generation count and the number of offspring for any given network.

2. Another issue worth exploring is the scalability of VRM_MOEA, which affects the performance of the framework. As we can see in Section 5.6.2.6, the computational time of VRM_MOEA increases linearly with the size of network. We can predict that if the network size is large, i.e. 100 nodes, the computational time will be long, e.g. 5 minutes by extrapolating the results where the generation count is 20000 (as given in Figure 49). Hence, the algorithm efficiency may need to be improved.
3. We used a VRM_MOEA to solve the destination physical platform selection problem considering two conflicting objectives. However, VRM_MOEA can be applied to other applications with similar features. For example, in a target-tracking problem in wireless sensor networks [191], some sensors nodes are chosen to track the target based on collected information, e.g. proximity to the target, their existing workload and residual battery life. These factors can be considered as objectives and then the nodes with good quality can be identified by VRM_ MOEA mechanism in order to create a suitable self-organising cluster to perform the necessary task.
4. In the traditional model, ISPs usually own the physical resource and provide the service. In the business and network virtualization model, the traditional single role of ISPs is divided into two different ones: that of an infrastructure provider and a service provider. Infrastructure providers manage the physical network resources whilst a service provider leases the physical resources from one or several infrastructure providers to create virtual networks providing customized services. The one to many, many to one and many to many relationship between infrastructure providers and service providers may bring challenges to our energy management scheme as some negotiation are required between them. For example,

the service provider may not agree with virtual router migration as they need virtual router instances to be located near to their customers in order to provide adequate real-time services. Hence, the energy management scheme could be customized to the preference of service providers in future work.

5. One promising next generation network architecture to replace traditional IP-style packet-switch host-centric networking is Named Data Networking (NDN)/ Content–Centric Networking (CCN) which focuses on large-scale content distribution [194][195]. There are two main types of packet, i.e. data packets and interest packets. An interest packet contains the content that a customer is interested in and the router will send it to the data producer(s). Then, data packets with related information are sent back from the data producer(s) to the customer. One feature of NDN is that NDN routers can have massive-scale caches (e.g. terabyte-scale caching) which cache previously forwarded data packets [196][197]. These data packets can then be reused when a matching interest packet arrives, which provides a quick response to customer request(s). However, NDN architecture may bring challenges to our energy management scheme. For instance, if an NDN router instance is moved to another physical location, it may bring some delay when data packets in the cache of a remotely located VR instance are transmitted to customers.

REFERENCES

[1] Bolla, R.; Bruschi, R.; Davoli, Franco; Cucchietti, F., "Energy Efficiency in the Future Internet: A Survey of Existing Approaches and Trends in Energy-Aware Fixed Network Infrastructures," *IEEE Communications Surveys & Tutorials*, vol.13, no.2, pp.223-244, Second Quarter 2011.

[2] Bianzino, A.P.; Chaudet, C.; Rossi, D.; Rougier, J., "A Survey of Green Networking Research," *IEEE Communications Surveys & Tutorials*, vol.14, no.1, pp.3-20, First Quarter 2012.

[3] Zhang,Y.; Chowdhury, P.; Tornatore, M.; Mukherjee, B., "Energy Efficiency in Telecom Optical Networks," *IEEE Communications Surveys & Tutorials*, vol.12, no.4, pp.441-458, Fourth Quarter 2010.

[4] Yamanaka, N.; Takeshita, H.; Okamoto, S.; Hirao, A.; Sato, T., "Creating Future Energy Efficient Network Services through Optical Technologies," in *2013 First International Symposium on Computing and Networking (CANDAR)*, pp.8-14, 4-6 Dec. 2013

[5] Lange, C.; Kosiankowski, D.; von Hugo, D.; Gladisch, A., "Analysis of energy consumption in carrier networks," in *2014 International Conference on Optical Network Design and Modeling*, pp.96-101, 19-22 May 2014.

[6] Lange, C.; Kosiankowski, D.; Weidmann, R.; Gladisch, A., "Energy Consumption of Telecommunication Networks and Related Improvement Options," *IEEE Journal of Selected Topics in Quantum Electronics*, vol.17, no.2, pp.285-295, March-April 2011.

[7] Tucker, R.S., "Energy consumption in telecommunications," in *2012 IEEE Optical Interconnects Conference*, pp.1-2, May 2012.

[8] Cisco® Visual Networking Index, "Cisco VNI: Forecast and Methodology, 2012-2017."[online].Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.html[Accessed May/10/2014].

[9] Cisco® Visual Networking Index, "Global Mobile Data Traffic Forecast Update, 2013–2018."[online]. Available:http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html [Accessed August/10/2014].

[10] Baliga, J.; Ayre, R.; Hinton, K.; Tucker, R.S., "Photonic Switching and the Energy Bottleneck," in *2007 Photonics in Switching*, pp.125-126, August 2007.

[11] Cisco CRS Multishelf System Overview; [online]. Available:http://www.cisco.com/c/en/us/td/docs/routers/crs/crs1/mss/16_slot_fc/system_description/reference/guide/sysdesc/msssd1.html[Accessed August/10/2014].

[12] Neilson, D.T., "Photonics for Switching and Routing," *IEEE Journal of Selected Topics in Quantum Electronics*, vol.12, no.4, pp.669-678, July-August 2006.

[13] Bohr, M., "A 30 Year Retrospective on Dennard's MOSFET Scaling Paper," *IEEE Solid-State Circuits Society Newsletter*, vol.12, no.1, pp.11-13, winter 2007.

[14] UK department of energy & climate change report, "Energy Trends: September 2013,"[online].Available:https://www.gov.uk/government/publications/energy- -september-2013[Accessed August/10/2014].

[15] British Telecom, "BT Group plc Annual Report& Form 20-F 2013,"[online]. Available:http://www.btplc.com/Sharesandperformance/Annualreportandreview/pdf/2013_BT_Annual_Report_smart.pdf [Accessed August/10/2014].

[16] Global e-Sustainibility Initiative, "SMART 2020: Enabling the Low Carbon Economy in the Information Age,"[online].Available: http://www.smart2020.org/_assets/files/02_Smart2020Report.pdf[Accessed August/10/2014].

[17] Shakun, J. D.; Clark, P. U.; He, Feng *et al.*"Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation," *Nature,* vol.484, pp. 49-54, April 2012.

[18] Peter M. Vitousek; "Beyond Global Warming: Ecology and Global Change," *Ecology*, vol. 75, pp.1861–1876. October 1994.

[19] European Commission, "An ROAD Map for moving to a competitive low carbon economy in 2050", [online]. Available: http://www.cbss.org/wp-content/uploads/2012/12/EU-Low-Carbon-Road-Map-2050.pdf[Accessed August/10/2014].

[20] Brill, K., "2005-2010 Heat Density Trends in Data Processing, Computer Systems, and Telecommunications Equipment,"[online]. Available: assets/contents/files/2005_2010_HeatDensity.pdf[Accessed August/10/2014]

[21] Hernon, D.; Salamon, T.; Kempers, R., et.al. , "Thermal Management: Enabling Enhanced Functionality and Reduced Carbon Foot Print," *Bell Labs Technical Journal*, vol.14, issue 13, pp.7-19, November 2009.

[22] Feng, D.; Jiang,C.; Lim,G.; Cimini, L.J., Jr.; Feng,G.; Li, G.Y., "A survey of energy-efficient wireless communications," *IEEE Communications Surveys & Tutorials*, vol.15, no.1, pp.167-178, First Quarter 2013.

[23] Gupta, M. ;Singh, S., "Greening of the Internet," in *proceedings of 2003 SIGCOMM*, pp. 19-26, August 2003.

[24] Baliga, J.; Hinton, K.; Tucker, R.S., "Energy Consumption of the Internet," in *2007 Joint International Conference on Optical Internet and 32nd Australian Conference on Optical Fibre Technology*, pp.1-3, June 2007.

[25] Shen ,G.; Tucker, R.S., "Energy-Minimized Design for IP Over WDM Networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol.1, no.1, pp.176-186, June 2009.

[26] Tucker, R.S.; Parthiban, R.; Baliga, J.; Hinton, K.; Ayre, R.W.A.; Sorin, W.V., "Evolution of WDM Optical IP Networks: A Cost and Energy Perspective," *Journal of Lightwave Technology*, vol.27, no.3, pp.243-252, February 2009.

[27] Wang, L.; Lu, R.; Li, Q.; Zheng, X.; Zhang, H., "Energy efficient design for multi-shelf IP over WDM networks," in *2011 IEEE Conference on Computer Communications Workshops*, pp.349-354, April 2011.

[28] Nedevschi, S.; Popa, L.; Iannaccone, G.; Ratnasamy, S. ; Wetherall, D., "Reducing network energy consumption via sleeping and rate-adaptation," in *Proceeding of 2008 5th USENIX Symposium on Networked Systems Design and Implementation*, pp. 323-336, April 2008.

[29] Chiaraviglio, L.; Mellia, M.; Neri,F. ,"Energy-aware networks: Reducing power consumption by switching off network elements," in *FEDERICA-Phosphorus tutorial and workshop*,2008.

[30] Chiaraviglio, L.; Mellia, M.; Neri, F., "Reducing Power Consumption in Backbone Networks," in *2009 IEEE International Conference on Communications*, pp.1-6, June 2009.

[31] Chiaraviglio, L.; Mellia, M.; Neri, F., "Energy-Aware Backbone Networks: A Case Study," in *2009 IEEE International Conference on Communications Workshops*, pp.1-5, June 2009.

[32] Bianzino, A.P.; Chiaraviglio, L.; Mellia M.; Rougier, Jean-Louis, "GRiDA: GReen Distributed Algorithm for energy-efficient IP backbone networks, "*Computer Networks*, vol.56, no. 14, pp. 3219-3232, September 2012.

[33] Chabarek, J.; Sommers, J.; Barford, P.; Estan, C.; Tsiang, D.; Wright, S., "Power Awareness in Network Design and Routing," in *IEEE 27th Conference on Computer Communications*, April 2008.

[34] Yunlei Lui; Gangxiang Shen; Weidong Shao, "Design for energy-efficient IP over WDM networks with joint lightpath bypass and router-card sleeping strategies," Optical Communications and Networking, IEEE/OSA Journal of , vol.5, no.11, pp.1122-1138, Nov. 2013

[35] Cianfrani, A.; Eramo, V.; Listanti, M.; Marazza, M.; Vittorini, E., "An Energy Saving Routing Algorithm for a Green OSPF Protocol," in *2010 IEEE Conference on Computer Communications Workshops*, pp.1-5, March 2010.

[36] F. Idzikowski, E. Bonetto, L. Chiaraviglio, A. Cianfrani, A. Coiro, R. Duque, F.Jimenez, E. Le Rouzic, F. Musumeci, W. Van Heddeghem, J. L ´ opez Vizca ´ ´ıno, and Y. Ye. "TREND in Energy-Aware Adaptive Routing Solutions". *IEEE Communications Magazine*, pp. 94–104, November 2013.

[37] Cisco white paper, "Converge IP and DWDM Layers in the Core Network, " [online].Available: http://www.webtorials.com/main/resource/papers/cisco/paper114 /ConvergeIPandDWDMLayersintheCoreNetwork.pdf[Accessed August/10/2014].

[38] Gencata, A; Mukherjee, B., "Virtual-topology adaptation for WDM mesh networks under dynamic traffic," *IEEE/ACM Transactions on Networking,* vol.11, no.2, pp.236-247, Apr 2003.

[39] Hinton, K.; Baliga, J.; Feng, M.Z.; Ayre, R.W.A.; Tucker, R.S., "Power consumption and energy efficiency in the internet," *IEEE Network*, vol.25, no.2, pp.6,12, March-April 2011

[40] Uhlig, S.; Quoitin, B.; Lepropre, J.; Balon, S., "Providing public intradomain traffic matrices to the research community, "*ACM SIGCOMM Computer Communication Review* vol.36, no. 1, pp.83-86, January 2006.

[41] Abilene Network Traffic Statistics. [Online]. Available: http://noc.net.internet2.edu/ [Accessed August/10/2014]

[42] Bolla, R.; Bruschi, R.; Cianfrani, A.; Listanti, M., "Enabling backbone networks to sleep," *IEEE Network*, vol.25, no.2, pp.26-31, March-April 2011.

[43] Agrawal, M.; Bailey, S. R.; Greenberg, A.; Pastor, J.; Sebos, P.; Seshan, S.; Van Der Merwe, K.; Yates, J., "Router Farm: Towards a dynamic, manageable network edge," In *Proceedings of2006 SIGCOMM workshop*, pp. 5-10,August 2006.

[44] Wang, Y.; Keller, E.; Biskeborn, B.; Van der Merwe, J.; Rexford, J., "Virtual routers on the move: live router migration as a network-management primitive," in

ACM SIGCOMM Computer Communication Review. vol. 38,no. 4, pp. 231-242, January2008.

[45] CISCO white paper, "IP over DWDM: Empower Your Next-Generation Optical Network," [online].Available:http://www.cisco.com/c/en/us/products/collateral/optical-networking/ons-15454-series-multiservice-provisioning-platforms/white_paper_c11-601399.html[Accessed August/10/2014]

[46] Dixit S., *IP Over WDM: Building the Next-generation Optical Internet,* Wiley-Blackwell, April 2003.

[47] ITU-T, "*G.803 : Architecture of transport networks based on the synchronous digital hierarchy*"[online].Available:http://www.ntc-sss.ru/nali_pb_files/g.803/G.803.pdf [Accessed August/10/2014]

[48] Minzer, S.E., "Broadband ISDN and asynchronous transfer mode," *IEEE Communications Magazine,* vol. 27, no. 9, pp. 17-24, September 1989.

[49] Jun Zheng, Hussein T. Mouftah, *Optical WDM networks: concepts and design principles*, Wiley-IEEE Press, August 2004.

[50] Cisco datasheet, "Cisco IP over DWDM Solution: Transport for the Approaching Zettabyte Era," [online] Available: http://www.cisco.com/en/US/prod/collateral/routers/ps5763/Cisco_IPoDWDM_bro.pdf [Accessed August/10/2014]

[51] Tibuleac, S., "ROADM network design issues," in *2009 Optical Fiber* Communication Conference, pp.1-48, March 2009.

[52] Heavy Reading white paper, "The need for next generation ROADM networks,"[online]Available:http://www.jdsu.com/productliterature/NG_ROADM_WP_final2 .pdf [Accessed August/10/2014]

[53] Strasser,T.A.; Wagener,J.L., "Wavelength-Selective Switches for ROADM Applications," *IEEE Journal of Selected Topics in Quantum Electronics,* vol.16, no.5, pp.1150-1157, September 2010.

[54] J. Homa and K. Bala, "ROADM architectures and their enabling WSS technology," *IEEE Communications Magazine*,vol.46, no.7, pp.150-154, July 2008.

[55] Sabella, R.; Testa, F.; Iovanna, P.; Bottari, G., "Flexible packet-optical integration in the cloud age: Challenges and opportunities for network delayering," *IEEE Communications Magazine*, vol.52, no.1, pp.35-43, January2014.

[56] Yongcheng Li; Li Gao; Gangxiang Shen; Limei Peng, "Impact of ROADM colorless, directionless, and contentionless (CDC) features on optical network performance," *IEEE/OSA Journal of Optical Communications and Networking*, vol.4, no.11, pp.58-67, November2012.

[57] Gangxiang Shen; Yongcheng Li; Limei Peng, "How much can colorless, directionless and contentionless (CDC) of ROADM help dynamic lightpath provisioning?," in *2012 Optical Fibre Communication Conference and Exposition and the National Fibre Optic Engineers Conference*, pp.1-3, March 2012.

[58] Gringeri, S.; Basch, B.; Shukla, V.; Egorov, R.; Xia, T.J., "Flexible architectures for optical transport nodes and networks," *IEEE Communications Magazine*, vol.48, no.7, pp.40-50, July 2010.

[59] Turkcu, O.; Subramaniam, S., "Performance of Optical Networks with Limited Reconfigurability," *IEEE/ACM Transactions on Networking*, vol.17, no.6, pp.2002-2013, December 2009.

[60] Advanced Configuration and Power Interface Specification [online] Available:http://www.acpi.info/DOWNLOADS/ACPIspec50[Accessed August/10/2014].

[61] Christensen, K.; Nordman, B.; Brown, R., "Power management in networked devices," *Compute*, vol.37, no.8, pp.91-93, August2004.

[62] Ceuppens, L.; Sardella, A.; Kharitonov, D., "Power Saving Strategies and Technologies in Network Equipment Opportunities and Challenges, Risk and Rewards," in *2008International Symposium on Applications and the Internet*, pp.381-384, July 2008.

[63] Roberts, L.G., "A radical new router," *IEEE Spectrum*, vol.46, no.7, pp.34-39, July 2009.

[64] Baldi, M.; Ofek, Y., "Time for a "Greener" Internet," in *IEEE International Conference on Communications Workshops,* pp.1-6, June2009.

[65] Chengchen Hu; Chunming Wu; Wei Xiong; Binqiang Wang; Jiangxing Wu; Ming Jiang, "On the design of green reconfigurable router toward energy efficient internet," *IEEE Communications Magazine*, vol.49, no.6, pp.83-87, June 2011.

[66] Yi Kai; Bin Liu; Jianyuan Lu,"Green Router: Power-Efficient Router Design, "in *Proceedings of International Conference on Computer Science and Information TechnologyAdvances in Intelligent Systems and Computing*, vol. 255, pp.197-204, September 2013.

[67] Barroso,L. A.; Holzle, U., "The Case for Energy-Proportional Computing," *Computer*, vol.40, no.12, pp.33-37, December2007.

[68] Tucker, R.S., "Scalability and Energy Consumption of Optical and Electronic Packet Switching," *Journal of Lightwave Technology*, vol.29, no.16, pp.2410-2421, August 2011.

[69] Krishnamoorthy, A.V.; Goossen, K.W.; Jan, W.; Xuezhe Zheng; Ho, R.; Guoliang Li; Rozier, R.; Liu, F.; Patil, D.; Lexau, J.; Schwetman, H.; Dazeng Feng; Asghari, M.; Pinguet, T.; Cunningham, J.E., "Progress in Low-Power Switched Optical

Interconnects," *IEEE Journal of Selected Topics in Quantum Electronics*,vol.17, no.2, pp.357-376, March-April 2011.

[70] Mishra, A.; Tripathi,A.K., "Energy efficient voltage scheduling for multi-core processors with software controlled dynamic voltage scaling,"*Applied Mathematical Modelling*, January 2014.

[71] David, H.; Fallin, C.; Gorbatov, E.; Hanebutte, U.R.; Mutlu, O., "Memory power management via dynamic voltage/frequency scaling, "in *Proceedings of the 8th ACM international conference on Autonomic computing* , pp.31-40, June 2011.

[72] Gunaratne, C.; Christensen, K.; Nordman, B.; Suen, S., "Reducing the Energy Consumption of Ethernet with Adaptive Link Rate (ALR)," *IEEE Transactions on Computers*, vol.57, no.4,pp.448-461, April 2008.

[73] Gupta, M.; Grover, S.; Singh, S., "A feasibility study for power management in LAN switches," in *Proceedings of the 12th IEEE International Conference on Network Protocols*, pp.361-371, October2004.

[74] Gupta, M.; Singh, S., "Using Low-Power Modes for Energy Conservation in Ethernet LANs," in*26th IEEE International Conference on Computer Communications*,pp.2451-2455, May 2007.

[75] Orgerie,A.C.; Lefèvre,L., "ERIDIS: Energy-Efficient Reservation Infrastructure for Large-Scale Distributed Systems, "*Parallel Processing Letters*, vol. 21, pp.133-154, June 2011.

[76] IEEE, "IEEE P802.3az Energy Efficient Ethernet Task Force," [online] Available: http://www. ieee802.org/3/az/index.html [Accessed August/10/2014]

[77] Christensen, K.; Reviriego, P.; Nordman, B.; Bennett, M.; Mostowfi, M.; Maestro, J.A., "IEEE 802.3az: the road to Energy Efficient Ethernet," *IEEE Communications Magazine*, vol.48, no.11, pp.50-56, November 2010.

[78] D-LINK news, "First company to offer ‘green’ Ethernet technology for network connectivity, embrace energy saving initiatives"[online] Available:http://www.dlinkgreen.com/press.asp?pressrelease_id=6[Accessed August/10/2014]

[79] Kelly, T.; Adolph, M., "ITU-T initiatives on climate change," *IEEE Communications Magazine*, vol. 46, pp. 108-114, October 2008.

[80] Sustainability in Telecom: Energy and Protection Committee (STEP),[online] Available: http://www.atis.org/STEP/[Accessed August/10/2014]

[81] ATIS Develops Telecommunications Energy Efficiency Ratio,[online] Available: http://www.atis.org/PRESS/pressreleases2009/031909.htm[Accessed August/10/2014]

[82] Yoo, S.J.B., "Energy Efficiency in the Future Internet: The Role of Optical Packet Switching and Optical-Label Switching," *IEEE Journal of Selected Topics in Quantum Electronics*, vol.17, no.2, pp.406-418, March-April 2011.

[83] Eilenberger, G.J.; Bunse, S.; Dembeck, L.; Gebhard, Ul.;Ilchmann, F.; Lautenschlaeger, W.; Milbrandt, J., "Energy-efficient transport for the future Internet," *Bell Labs Technical Journal* , vol.15, no.2, pp.147-167, September 2010.

[84] Malkamaki, T.; Ovaska, S.J., "Data centres and energy balance in Finland," in *2012 International Green Computing Conference*, pp.1-6, June 2012.

[85] Tech Titans Building Boom, [online] Available: http://spectrum.ieee.org/green-tech/buildings/tech-titans-building-boom[Accessed August/10/2014]

[86] Bianzino, A.P.; Chaudet, C.; Larroca, F.; Rossi, D.; Rougier, J., "Energy-aware routing: A reality check," in *2010 IEEE GLOBECOM Workshops,* pp.1422-1427, December 2010.

[87] OSPF Extensions for MPLS Green Traffic Engineering, [online] Available: http://tools.ietf.org/html/draft-li-ospf-ext-green-te-00[Accessed August/10/2014]

[88] Thompson, K.; Miller, G.J.; Wilder, R., "Wide-area Internet traffic patterns and characteristics," *IEEE Network*, vol.11, no.6, pp.10-23, Nov/Dec 1997.

[89] Kihl, M.; Odling, P.; Lagerstedt, C.; Aurelius, A., "Traffic analysis and characterization of Internet user behavior," in *2010 International Congress on Ultra-Modern Telecommunications and Control Systems and Workshops*, pp.224-231, October 2010.

[90] Basu, A.; Riecke,J., "Stability Issues in OSPF Routing," in *proceeding of 2001 SIGCOMM*, vol.31, no. 4, pp.225-236,October 2001.

[91] RFC 1583OSPF Version 2[online] Available: http://www.ietf.org/rfc/rfc1583.txt[Accessed August/10/2014]

[92] Idzikowski,F.; Orlowski,S.; Raack, C.; Woesner,H.; Wolisz,A., "Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios," in *Conference on Optical Network Design and Modeling*, pp. 1-6, February2010.

[93] Zhang,Y.; Tornatore,M.; Chowdhury,P.; Mukherjee, B., "Time-Aware Energy Conservation in IP-over-WDM Networks," in *2010 Photonics in Switching*, July 2010.

[94] Fisher,W.; Suchara,M.; Rexford,J., "Greening backbone networks: reducing energy consumption by shutting off cables in bundled links," in *2010 Green Networking,* pp. 29-34, 2010.

[95] Bonetto, E.; Chiaraviglio, L.; Cuda, D.; Idzikowski, F.; Neri, F., "Exploiting traffic dynamics in Power-Aware Logical Topology Design," in *2011 37th European Conference and Exhibition on Optical Communication*, pp.1-3, September 2011.

[96] Yayimli, A.; Cavdar, C.;, "Energy-aware virtual topology reconfiguration under dynamic traffic," in *2012 14th International Conference on Transparent Optical Networks* , pp.1-4, July 2012.

[97] Muhammad, A.; Monti, P.; Cerutti, I.; Wosinska, L.; Castoldi, Piero; Tzanakaki, A., "Energy-Efficient WDM Network Planning with Dedicated Protection Resources in Sleep Mode," in *2010 IEEE Global Telecommunications Conference*, pp.1-5, December 2010.

[98] Mosharaf, N.M.; Chowdhury,K.; Boutaba,R., "A survey of network virtualization, Computer Networks, "*Computer Networks*, vol. 54, no.5, pp. 862-876, April 2010

[99] Chowdhury, N.M.M.K.; Boutaba, R., "Network virtualization: state of the art and research challenges," *IEEE Communications Magazine*, vol.47, no.7, pp.20-26, July 2009.

[100] Yu, M.; Yi, Y.; Rexford, J.; Chiang, M., "Rethinking Virtual Network Embedding: Substrate Support for Path Splitting and Migration," *ACM SIGCOMM Computer Communication Review*, vol. 38 no. 2, pp. 17-29, April 2008.

[101] Yong Zhu; Ammar, M., "Algorithms for Assigning Substrate Network Resources to Virtual Network Components," in *Proceedings25th IEEE International Conference on Computer Communications*, pp.1-12, April 2006 .

[102] Hongyan Cui; Wenjun Gao; Jiang Liu; Yunjie Liu, "A virtual network embedding algorithm based on virtual topology connection feature," in *16th International Symposium on Wireless Personal Multimedia Communications*, pp.1,5, June 2013.

[103] Report to congress on server and data centre energy efficiency public law 109-431.[online] available: http://www.energystar.gov/index.cfm?c=prod_development.server_efficiency_study[Accessed August/10/2014]

[104] RFC 2917 : A Core MPLS IP VPN Architecture, .[online] available: http://tools.ietf.org/html/rfc2917 [Accessed August/10/2014]

[105] Clark, C.; Fraser, K.; Hand, S., et.al. "Live Migration of Virtual Machines, "in *Proceedings of the 2nd conference on Symposium on Networked Systems Design &Implementation*, vol. 2, pp. 273-286, 2005.

[106] Haikun Liu; Hai Jin; Xiaofei Liao; Chen Yu; Cheng-Zhong Xu, "Live Virtual Machine Migration via Asynchronous Replication and State Synchronization," *IEEE Transactions on Parallel and Distributed Systems*, vol.22, no.12, pp.1986-1999, December 2011.

[107] Lo, S.; Ammar, M.; Zegura, E., "Design and analysis of schedules for virtual network migration," in *2013IFIP Networking Conference*, pp.1-9, May 2013.

[108] Keller, E.; Ghorbani, S.; Caesar, M.; Rexford, J., "Live migration of an entire network (and its hosts), " in *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, pp.109-114, October 2012.

[109] Heddeghem, W.V.; Idzikowski, F.; Vereecken, W.; Colle, D.; Pickavet, M.; Demeester, P, "Power consumption modelling in optical multiplayer networks," *Photonic Network Communications*, vol. 24, pp 86-102, October 2012.

[110] JayantBaliga, Robert Ayre, Kerry Hinton, Wayne V. Sorin, and Rodney S. Tucker, "Energy consumption in optical IP networks," *Journal of Lightwave Technology*, vol. 27, pp. 2391-2403, 2009.

[111] Kilper, D.C.; Atkinson, G.; Korotky, S.K.; Goyal, S.; Vetter, P.; Suvakovic, D.; Blume, O., "Power Trends in Communication Networks,", *IEEE Journal of Selected Topics in Quantum Electronics*, vol.17, no.2, pp.275-284, March-April 2011.

[112] CICSO cloud service router. [online] available:http://www.cisco.com/c/en/us/products/routers/cloud-services-router-1000v-series/index.html[Accessed August/10/2014]

[113] Juniper Logical router. [online] available: http://www.juniper.net/techpubs /software/junos/junos85/feature-guide-85/id-11139212.html[AccessedMay/10/2014]

[114] Bolla, R.; Bruschi, R.; Cianfrani, A.; Listanti, M., "Enabling backbone networks to sleep," *IEEE Network*, vol.25, no.2, pp.26-31, March-April 2011.

[115] Cohon, J. L., "Multiobjective Programming and Planning," New York: Academic Press,1978.

[116] Koski, J., "Multicriterion optimization in structural design,"NewDirections in Optimum Structural Design, pp. 483–503, Wiley,1984.

[117] Zitzler, E.; Thiele, L., "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE Transactions on Evolutionary Computation*, vol.3, no.4, pp.257-271, Nov 1999.

[118] Deb, K., Multi-Objective Optimization Using Evolutionary Algorithms. Wiley, Chichester, 2001.

[119] Ray, T.; Tai K.; Seow, K.C.,"Multiobjective Design Optimization by an Evolutionary Algorithm,"*Engineering Optimization*, vol. 33 pp. 399-424, 2001.

[120] Bäck, T.; Schwefel,H.P., "An overview of evolutionary algorithms for parameter optimization,"*Evolutionary Computation*, vol.3, no.1, pp.1-16, spring 1993.

[121] Eiben, A.E.; Schippers, C.A., "On evolutionary exploration and exploitation," *Fundamenta Informaticae*, vol.35, no.1-4, pp.35-50, 1998.

[122] Crepinsek, M.; Liu, S.H.; Mernik, M., "Exploration and Exploitation in Evolutionary Algorithms: A Survey," *ACM Computing Surveys*, vol.45, no.35, June 2013.

[123] Michalewicz, Z., Genetic Algorithms + Data Structures = Evolution Programs, Springer, January 1996.

[124] Grefenstette,J.; Gopal, R.; Rosmaita, B.; Gucht, D.V.,"Genetic algorithms for the travelling salesman problem," in *Proceedings of the first International Conference on Genetic Algorithms and their Applications*, pp.160-168, 1985.

[125] Goh, K.S.; Lim, A.; Rodrigues, B. "Sexual Selection for Genetic Algorithms,"*Artificial Intelligence Review*, vol.19, pp.123-152, 2003.

[126] Ponsich, A.; Jaimes, A.L.; Coello, C.A.C., "A Survey on Multiobjective Evolutionary Algorithms for the Solution of the Portfolio Optimization Problem and Other Finance and Economics Applications," *IEEE Transactions on Evolutionary Computation*, vol.17, no.3, pp.321-344, June 2013.

[127] Mukhopadhyay, A.; Maulik, U.; Bandyopadhyay, S.; Coello, C.A.C., "A Survey of Multiobjective Evolutionary Algorithms for Data Mining: Part I," *IEEE Transactions on Evolutionary Computation*, vol.18, no.1, pp.4-19, February 2014.

[128] Jourdan, L.; Corne, D.; Savic, D.; Walters, G., "Hybridising rule induction and multi-objective evolutionary search for optimising water distribution systems," in *Fourth International Conference on Hybrid Intelligent Systems*, pp.434-439, December 2004.

[129] Ehrgott, M.,Multicriteria Optimization, Springer-Verlag Berlin, 2000.

[130] Hajela, P.;Lin, C., "Genetic search strategies in multicriterion optimal design,"*Structural Optimization*, vol.4, no.2, pp.99-107,June 1992.

[131] Back, T.; Hammel, U.; Schwefel, H.-P., "Evolutionary computation: comments on the history and current state," *IEEE Transactions on Evolutionary Computation*, vol.1, no.1, pp.3-17, April 1997.

[132] Coello, C.A.C., "Evolutionary multi-objective optimization: a historical view of the field," *IEEE Computational Intelligence Magazine*, vol.1, no.1, pp. 28-36, February 2006.

[133] Srinivas,N.;Deb,K.,"Multiobjective optimization using nondominated sorting in genetic algorithms,"*Evolutionary Computation*, vol. 2, no. 3, pp. 221–248, Fall 1994.

[134] Horn,J.;Nafpliotis,N.;Goldberg,D.E.,"A niched pareto genetic algorithm for multiobjective optimization," In *Proceedings of the First IEEE Conference on Evolutionary Computation,* vol. 1, pp. 82–87, June 1994.

[135] Fonseca, C.M.;Fleming,P.J.,"Genetic algorithms for multiobjective Optimization: Formulation, discussion and generalization," In *Proceedings of the Fifth International Conference on Genetic Algorithms*, vol.93, pp.416-423, July 1993.

[136] Zitzler,E.;Laumanns,M.;Thiele,L., "SPEA2: Improving the strength Pareto evolutionary algorithm," In *proceedings of Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems*, pp. 95–100, 2002.

[137] Knowles,J.D.;Corne,D.W.,"Approximating the nondominated front using the Pareto archived evolution strategy,"*Evolutionary Computation*, vol.8, no.2, pp.149–172, 2000.

[138] Deb,K.;Agrawal,S.;Pratab,A.;Meyarivan,T.,"A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II," In *Proceedings of the Parallel Problem Solving from Nature VI Conference*, pp. 849–858, 2000.

[139] Coello, C.A.C.; Lamont, G.B.; Van Veldhuizen, D.A., Evolutionary Algorithms for Solving Multi-Objective Problems, 2nd edition, Springer, New York, 2007.

[140] Johannes, B.; Zitzler,E.,"HypE: An algorithm for fast hypervolume-based many-objective optimization,"*Evolutionary Computation*, vol.19, no.1, pp.45-76, 2011.

[141] Tang, L.; Wang,X., "A Hybrid Multiobjective Evolutionary Algorithm for Multiobjective Optimization Problems," *IEEE Transactions on Evolutionary Computation*, vol.17, no.1, pp.20-45, February 2013.

[142] Zhou,A.;Qu,B.;Li,H.; Zhao, S.;Suganthan, P.;Zhang,Q.,"Multiobjective evolutionary algorithms: A survey of the state of the art,"*Swarm and Evolutionary Computation*, vol.1, no.1, pp.32-49, March 2011.

[143] Von Lucken, C.; Baran, B.; Brizuela, C., "A survey on multi-objective evolutionary algorithms for many-objective problems,"*Computational Optimization and Applications*, pp.1-50,February2014.

[144] Zitzler, E.;Laumanns, M.;Bleuler, S.,"A tutorial on evolutionary Multiobjectiveoptimization,"*Lecture Notes in Economics and Mathematical System: Metaheuristics for Multiobjective Optimisation*,vol.535, pp.3-37, 2004.

[145] Kursawe, F.,"A variant of evolution strategies for vector optimization,"*Lecture Notes in Computer Science: Parallel Problem Solving from Nature*, vol, 496, pp.193–197, 1991.

[146] Silverman, B.W., Density Estimation for Statistics and Data Analysis. New York: Chapman and Hall, 1986.

[147] Multi-Objective Optimization Using Genetic Algorithms: A Tutorial

[148] Yu T.; Davis, L.;Baydar, C.; Roy, R., An introduction to evolutionary computation in practice. Springer, 2008.

[149] Herrera, F.; Lozano, M.; Verdegay, J.L., "Tacking real-coded genetic algorithms: Operators and tools for behavioural analysis," *Artificial Intelligence*,vol.12, pp.265-319, 1998.

[150] Herrera,F.; Lozano, M.; Sánchez, A.M., "A taxonomy for the crossover operator for real-coded genetic algorithms: An experimental study," *International Journal of Intelligent Systems,*vol.18, no.3, pp.309-338, February 2003.

[151] Eshelman, L.J.;Schaffer, J.D.,"Real-Coded Genetic Algorithms and Interval-Schemata, ",*Foundations of Genetic Algorithms 2*,pp. 187-202, 1992.

[152] Zitzler, E.; Thiele, L.,"Multiobjective optimization using evolutionary algorithms-a comparative case study,"*Lecture Notes in Computer Science: Parallel Problem Solving from Nature*, vol.1498, pp.292-301, 1998.

[153] Zitzler, E.; Deb. K.; Thiele, L., "Comparison of Multiobjective Evolutionary Algorithms: Empirical Results,"*Evolutionary Computation*, vol.8, no.2, pp.173-195, summer 2000.

[154] Robinson,S., Simulation – The practice of model development and use. Wiley, 2004.

[155] Law, A.M., *Simulation Modeling and Analysis*, fourth edition, McGraw-Hill, 2007.

[156] Schruben,L.,"Simulationmodeling with event graphs,"*Communications of the ACM*, vol.26, pp.957-963, November 1983.

[157] Khare,V.; Yao,X.; Deb, K., "Performance Scaling of Multi-objective Evolutionary Algorithms,"*Evolutionary Multi-Criterion OptimizationLecture Notes in Computer Science,*vol. 2632, pp 376-390, April 2003.

[158] Zitzler, E.; Thiele, L.; Laumanns, M.; Fonseca, C.M.; da Fonseca, V.G., "Performance assessment of multiobjective optimizers: an analysis and review," *IEEE Transactions on Evolutionary Computation*, vol.7, no.2, pp.117-132, April 2003.

[159] Goldberg, D.E., Genetic algorithms in search, optimization and machine learning.Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA,1989.

[160] Cisco white paper, "IEEE 802.3az EEE: build greener networks,"[online] available:http://www.cisco.com/c/dam/en/us/products/collateral/switches/catalyst-4500-series-switches/white_paper_c11-676336.pdf[Accessed August/10/2014]

[161] Gu, Y.; Liu, Y.; Towsley, D., "On integrating fluid models with packet simulation," in *Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol.4, pp.2856-2866, March 2004.

[162] Kiddle, C.; Simmonds, R.; Williamson, C.; Unger, B., "Hybrid packet/fluid flow network simulation," in *Proceedings. Seventeenth Workshop on Parallel and Distributed Simulation*, pp.143-152, June 2003.

[163] Cisco CRS Multi-shelf System Overview,[online] available: http://www.cisco.com/c/en/us/td/docs/routers/crs/crs1/mss/16_slot_fc/system_description/reference/guide/sysdesc/msssd1.html#wp1051350[Accessed August/10/2014]

[164] CRS-1 and 3 series,[online] available:http://www.cisco.com/c/en/us/products/routers/crs-1-multishelf-system/index.html[Accessed August/10/2014]

[165] Cisco CRS Modular Services Cards Data Sheet, [online] available:http://www.cisco.com/c/en/us/products/collateral/routers/carrier-routing-system/datasheet-c78-730791.html[Accessed August/10/2014]

[166] Cisco CRS 4-Port 10GE Tunable WDMPHY Interface Module,[online] available:http://www.cisco.com/c/en/us/products/collateral/routers/crs-1-8-slot-single-shelf-system/product_data_sheet0900aecd80395b82.html[Accessed August/10/2014]

[167] Cisco CRS-1 16-Slot Single-Shelf System, [online] available:http://www.cisco.com/c/en/us/products/collateral/routers/crs-1-16-slot-single-shelf-system/product_data_sheet09186a008022d5f3.html[Accessed August/10/2014]

[168] Cisco CRS-1 24-Slot Fabric-Card Chassis[online] available:http://www.cisco.com/c/en/us/products/collateral/routers/crs-1-16-slot-single-shelf-system/product_data_sheet0900aecd80340baa.html[Accessed August/10/2014]

[169] Vasudevan,R.; Morley Mao, Z.,"Reval: A Tool for Real-time Evaluation of DDoS Mitigation Strategies," USENIX 2006 Annual Technical Conference, 2006.

[170] Tangmunarunkit, H.;Govindan, R.;Jamin, S.;Shenker, S.;Willinger, W.,"Network topology generators: degree-based vs. structural,"in *Proceedings of the 2002*

Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp. 147-159, 2002.

[171] B.M.Waxman, "Routing of multipoint connections,"*IEEE Journal on Selected Areas in Communications*,vol,6, pp.1617-1622,1988.

[172] Zegura,E.;Calvert,K.L.;Donahoo, M. J.,"A Quantitative Comparison of raph-Based Models for Internet Topology,"*IEEE/ACM Transactions in Networking*, vol.5, pp.770-783, December1997.

[173] Calvert, K.L.; Doar, M.B.; Zegura, E.W., "Modeling Internet topology," *IEEE Communications Magazine*, vol.35, no.6, pp.160-163, June 1997.

[174] Doar, M.B., "A better model for generating test networks," in *Global Telecommunications Conference*, pp.86-93, Nov 1996.

[175] Faloutsos,M.;Faloutsos,P.;Faloutsos,C.,"On power-law relationships of the Internet topology,"in *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication* , pp.251-262,August 1999.

[176] Winick,J.;Jamin,S.,"Inet-3.0: Internet topology generator,"University of Michigan, Tech. Rep., 2002.

[177] Medina,A.;Matta,I.;Byers,J.,"On the origin of powerlaws in Internet topologies,"*ACM SIGCOMM Computer Communication Review*, vol. 30, no. 2, pp. 18–28, April 2000.

[178] Boston University, "BRITE: Boston University Representative InternetTopology Generator,"[online]available:http://www.cs.bu.edu/brite/[Accessed August/10/2014]

[179] J. Winick and S. Jamin,"Inet-3.0: Internet topology generator,"Tech.Rep., 2002.

[180] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A First-Principles Approach to Understanding the Internet's Router-level Topology", in *Proceedings of the 2004*

conference on Applications, technologies, architectures, and protocols for computer communications, pp.3-14, 2004.

[181] Y. Zhang, M. Roughan, C. Lund, and D. Donoho,"An Information-Theoretic Approach to Traffic Matrix Estimation,"in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pp.301-312, 2003.

[182] InternetTraffic Archive, [online] available: http://www.acm.org/sigcomm/ITA/[Accessed August/10/2014]

[183] University of Waikato, The DAG project, [online] available: http://dag.cs.waikato.ac.nz/[Accessed August/10/2014]

[184] University of Oregon Route Views Project, [online] available:http://www.routeviews.org/[Accessed August/10/2014]

[185] Nanog Looking Glass Sites, [online] availablehttp://www.nanog.org/lookingglass[Accessed August/10/2014]

[186] AbileneTM,[online] available:http://www.cs.utexas.edu/~yzhang/research/AbileneTM/[Accessed August/10/2014]

[187] The Abilene Network, [online] available: http://www.internet2.edu/pubs/200502-IS-AN.pdf[Accessed August/10/2014]

[188] Grefenstette, J.J., "Optimization of Control Parameters for Genetic Algorithms," *IEEE Transactions on Systems, Man and Cybernetics,* vol.16, no.1, pp.122-128, Jan. 1986

[189] Eiben,A.E;Hinterding,R.;Michalewicz,Z.,"Parameter Control in Evolutionary Algorithms,"*IEEE Transactions on Evolutionary Computation*, pp.124–141, 1999.

[190] Smit, S. K.; Eiben, A.E., "Comparing parameter tuning methods for evolutionary algorithms," *IEEE Congress on Evolutionary Computation*, pp.399-406, May 2009.

[191] Hamouda, Y.E.M.; Phillips, C., "Metadata-Based Adaptive Sampling for Energy-Efficient Collaborative Target Tracking in Wireless Sensor Networks," in *2010 IEEE 10th International Conference on Computer and Information Technology*, pp.313-320, July 2010.

[192] Computer Power Usage, https://secure.www.upenn.edu/computing/resources/category/hardware/article/computer-power-usage

[193] Ciaramella, E., "Wavelength Conversion and All-Optical Regeneration: Achievements and Open Issues," Lightwave Technology, Journal of , vol.30, no.4, pp.572,582, Feb.15, 2012

[194] Katsaros, K.; Wei Chai; Ning Wang; Pavlou, G.; Bontius, H.; Paolone, M., "Information-centric networking for machine-to-machine data delivery: a case study in smart grid applications," *IEEE Network*, vol.28, no.3, pp.58-64, May-June 2014

[195] Ohtani, M.; Tsukamoto, K.; Koizumi, Y.; Ohsaki, H.; Imase, M.; Hato, K.; Murayama, J., "VCCN: Virtual content-centric networking for realizing group-based communication,", in *2013 IEEE International Conference on Communications (ICC)*, pp.3476,3480, 9-13 June 2013

[196] Xiaoyan Hu; Jian Gong, "Distributed in-network cooperative caching," in *2012 IEEE 2nd International Conference on Cloud Computing and Intelligent Systems (CCIS)*, vol.02, pp.735,740, Oct. 30 2012-Nov. 1 2012

[197] Hao Wu; Jun Li; Tian Pan; Bin Liu, "A novel caching scheme for the backbone of Named data networking," in *2013 IEEE International Conference on Communications (ICC),* , pp.3634-3638, 9-13 June 2013

APPENDIX A: TRAFFIC ANANLYSIS

In this section, an analysis of publicly traffic data from the Abilene network is presented. Part 1 introduces the motivation for traffic analysis. Then some related work is described in Part 2. In Part 3, detailed information concerning the Abilene network is examined. Next, traffic modelling is described in Part 4 and finally Part 5 provides a brief summary.

Part 1 Motivation

The research objective is to propose an energy-efficient network by reconfiguring resources. The main concept of the framework is to combine infrastructure sleeping and Virtual Router Migration (VRM). One of the basic problems in the framework is to determine when to move the virtual router (VR) instances to remote destination physical platforms. This problem involves another problem – whether the network traffic load can be predicted. Hence, the characteristics of the traffic demand need to be explored.

In this section, we want to solve the question as follows: whether the traffic demand matrix at a similar time from one day to the next is related or not. To put it another way, is the traffic that flows between a given ingress / egress point similar at roughly the same time (given a certain observation window) over successive days? If this is the case then prediction mechanisms can be used to anticipate the traffic demand and reconfigure the network based on this. Some offline optimization techniques can be used. Conversely, if there is no significant correlation in the traffic level then only reactive mechanisms are viable.

Part 2 Related Work

Traffic prediction uses previous traffic demand to forecast future values. Traffic prediction techniques can be divided into two types: short range forecasting (several minutes or hours) or long range forecasting (several days or months). Traffic prediction is a useful tool in network management and can be applied in many areas, e.g. network performance evaluation, buffer management and network capacity planning [1] [2].

In order to predict the network traffic, it is important to appreciate the natural behaviour of Internet traffic. Internet traffic was firstly treated as a Poisson process or a Poisson related process which is a memory-less system and the inter-arrival times are exponentially distributed [3]. A key feature of the "Poisson-like" traffic is that when it is aggregated, the traffic becomes smoother and less bursty.

However, around 1993, researchers discovered that it is not suitable to use a Poisson process to represent the Internet traffic as the Internet traffic has different features from the traditional telephone traffic. Leland *et.al.* [4] found that the traffic showed self-similarity. Self-similarity indicates that the structure of the data is similar across many time scales (observation intervals are from 100 seconds down to 0.01 seconds) [7]. The paper summarized that the aggregation of the traffic intensifies the burstiness instead of smoothing it, which is different from "Poisson-like" traffic. In addition, Internet traffic has another characteristic called Long-Range Dependence (LRD) [6]. LRD indicates that the value at any time is non-negligibly positive correlated with all future values. After self-similarity and LRD concepts were proposed, research interest has shifted from the traditional memory-less and Poisson based modelling systems to long memory and bursty systems.

In mathematics, the degree of LRD and self-similarity are measured in term of the Hurst parameter (H) [8], which provides a good measure of traffic "burstiness". However, identifying and estimating H is not easy. Many approaches have been proposed for estimating H, e.g. time domain, frequency and wavelet methods. However, the results of different methods are often inconsistent [9]. Therefore, more effort is needed to characterize Internet traffic.

A common traffic prediction approach is to use time series models to fit the traffic, e.g. an Auto-Regressive Integrated Moving Average (ARIMA) [12] process. ARIMA encompasses several time series models: Auto-Regressive (AR), Moving Average (MA), and Auto-Regressive Moving Average (ARMA). An AR model uses the weighted sum of order q previous values (denoted by AR (p)) plus a random shock to predict the estimated value. On the other hand, an MA model uses the weighted sum of order q previous random shock values (denoted by MA(q)) plus a random shock to obtain the estimated value. ARMA combines AR and MA models, which predicts a value by adding the weighted sum of p previous values and q previous shocks. AR, MA and ARMA models all require the input data to be stationary, which has the same average and variance regardless of the time period of the data. However, real world data do not always fit the stationary requirement. Hence, an ARIMA model was developed in order to handle non-stationary time series. In ARIMA, the data are differenced in order to obtain a stationary data series.

In the wavelet domain, some researchers use wavelet analysis to predict the traffic since wavelets are a natural way to describe the multi-scale characteristic of self-similarity [13]. In addition, there exists a hybrid prediction method combining time domain and wavelet domain methods. [9] used the Wavelet Multi-resolution Analysis (WMA) and linear time series models to develop a methodology for building a long-term

traffic prediction model for any two neighbouring nodes in the backbone network. Furthermore, machine-learning methods can be used for traffic prediction, e.g. neural networks [14] [15].

Part 3 Network Traffic Data

In this section, we explore traffic modelling with the Abilene network data [16] using an ARIMA model. The remainder of the section is organised as follows. Initially, Part 3.1 introduces how the data is collected. Then, Part 3.2 show the traffic data is modelled and used for predictions. Finally, Part 3.3 provides a discussion.

Part 3.1 Data Description

The Abilene network is a high-performance backbone network created by the Internet2 community [17]. The clients are mostly universities and some corporate and affiliate institutions in the United States. There are 12 nodes and 15 links. Data has been collected in terms of 144 pairs of source and destination nodes every 5 minutes for 24discontinuous weeks. The traffic rate unit is kbps.

A few data points are missing from the dataset. There are usually two methods to handle the missing data, i.e. ignoring or estimating the missing data. As we use a linear model whose estimated output is predicted by using the sum of previous weighted values, the data needs to be complete. Hence, the missing data are estimated by averaging the values of the same time between the previous day and the following day.

Moreover, the data over 24 weeks is not continuous. The data were collected starting from Monday for most of weeks whilst some weekly data were collected starting from Saturday. This may have some impact on the model in terms of the weekly pattern. The

weekly pattern means that the traffic load is usually quieter on the weekends than that of weekdays. In order to avoid the impact of weekly pattern, we removed the data of weekend. The detail is described in Part 3.2.

Part 3.2 Traffic Modelling

There are 144 pairs of sources and destinations in the network. Four pairs of traffic are randomly selected to perform the analysis. These four pairs are called P1, P2, P3 and P4. The 24 weeks of traffic of 4 pairs are shown from Figure 3.1 to Figure 3.4. The first week data are shown from Figure 3.5 to Figure 3.8.

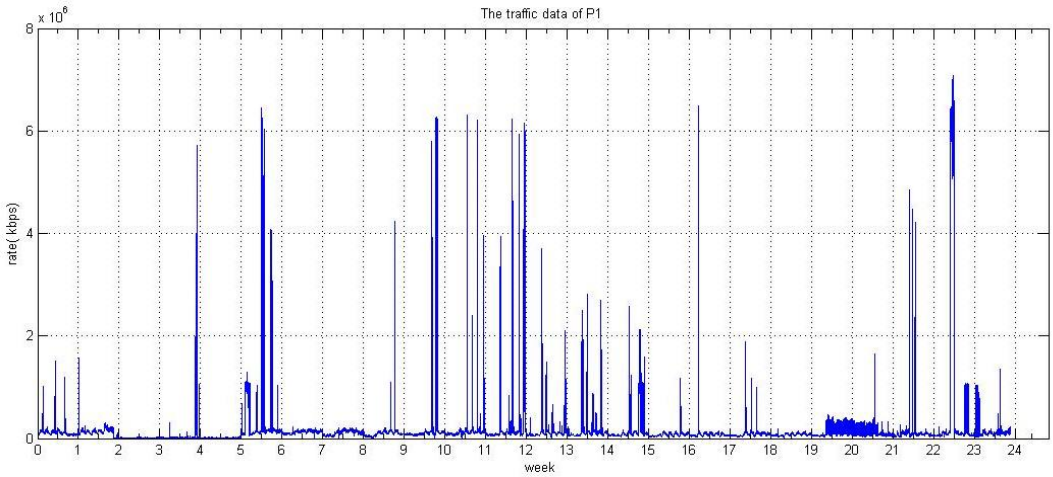

Figure 3.1 P1 Traffic Data

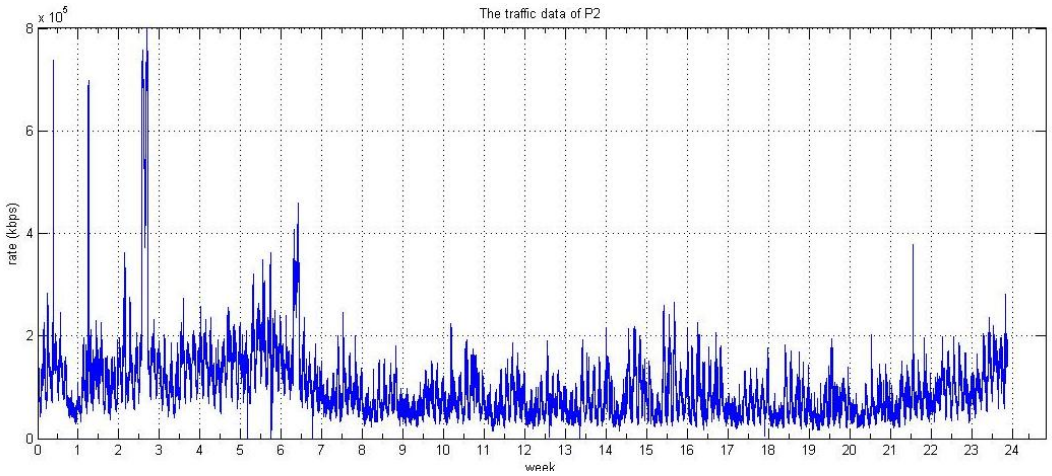

Figure 3.2 P2 Traffic Data

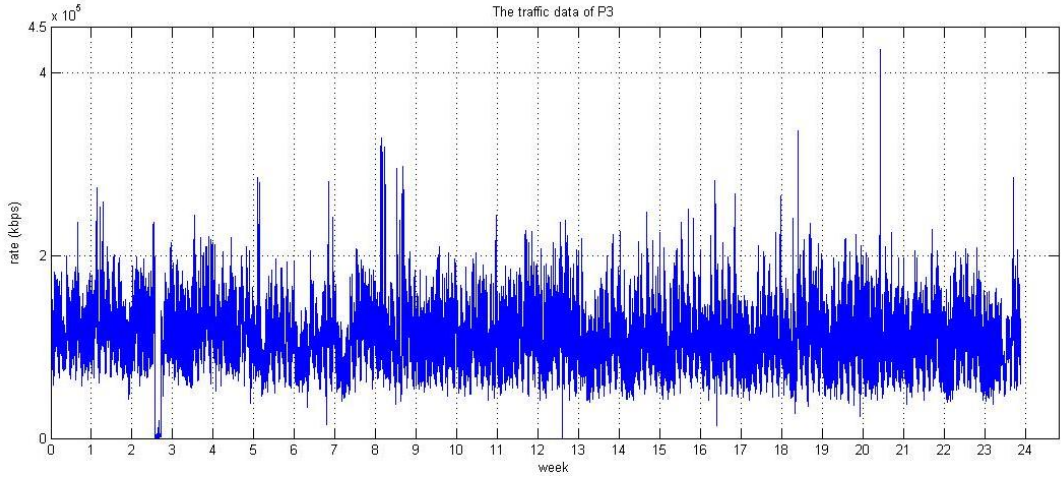

Figure 3.3 P3 Traffic Data

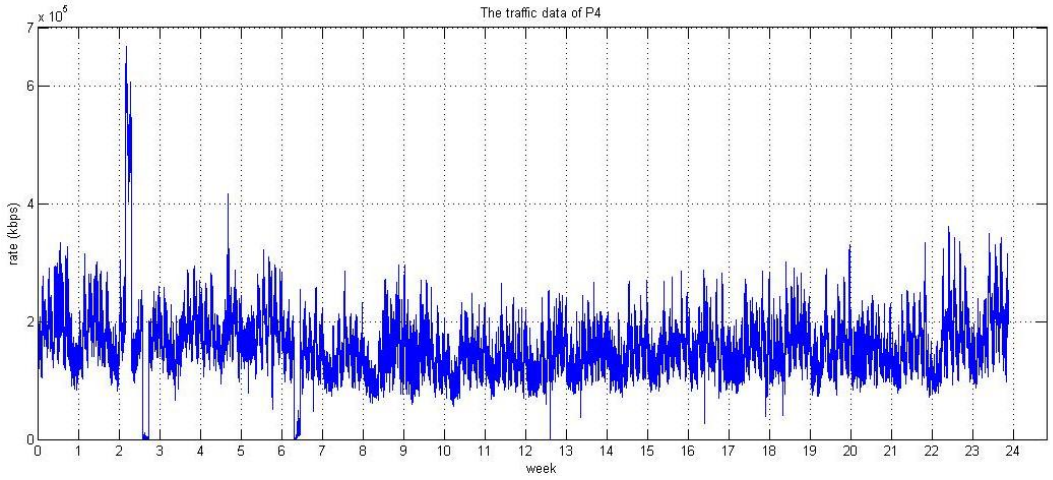

Figure 3.4 P4 Traffic Data

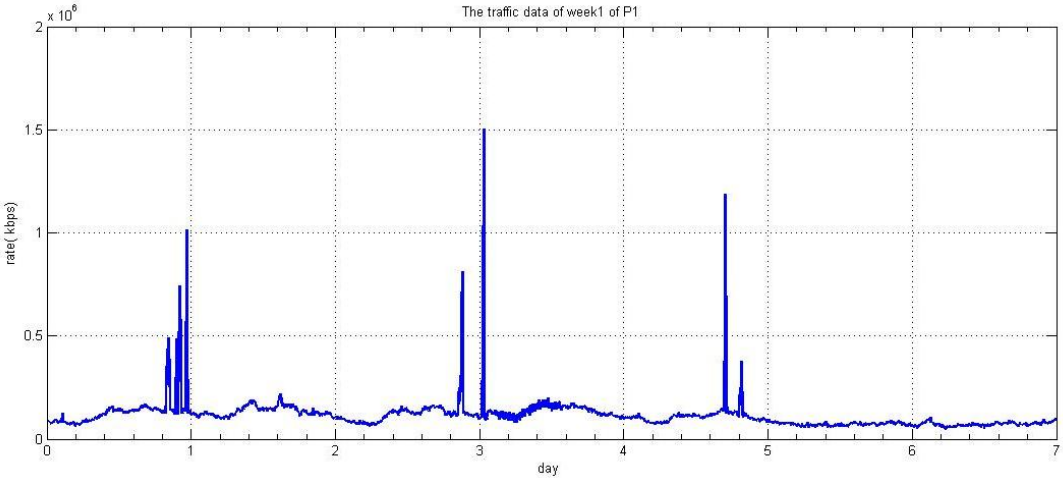

Figure 3.5 First Week Traffic Data of P1

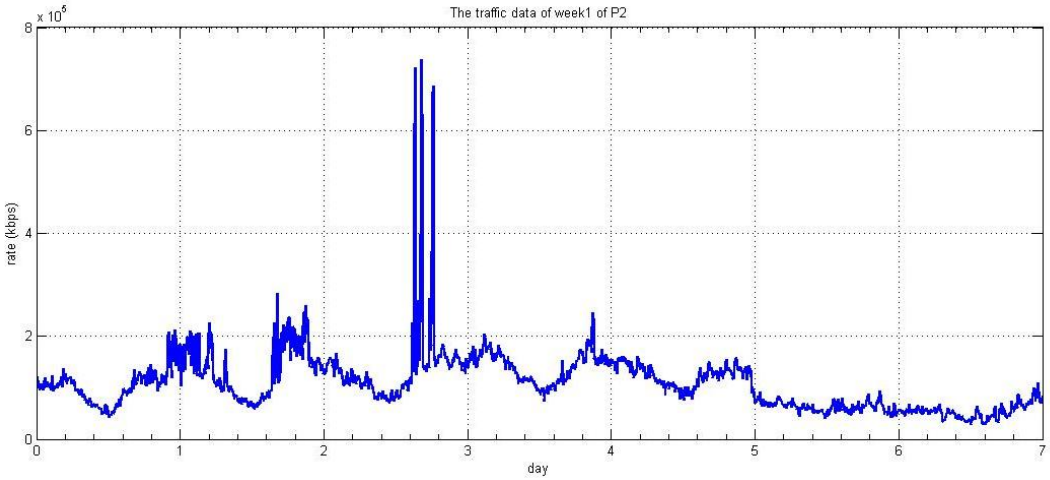

Figure 3.6 First Week Traffic Data of P2

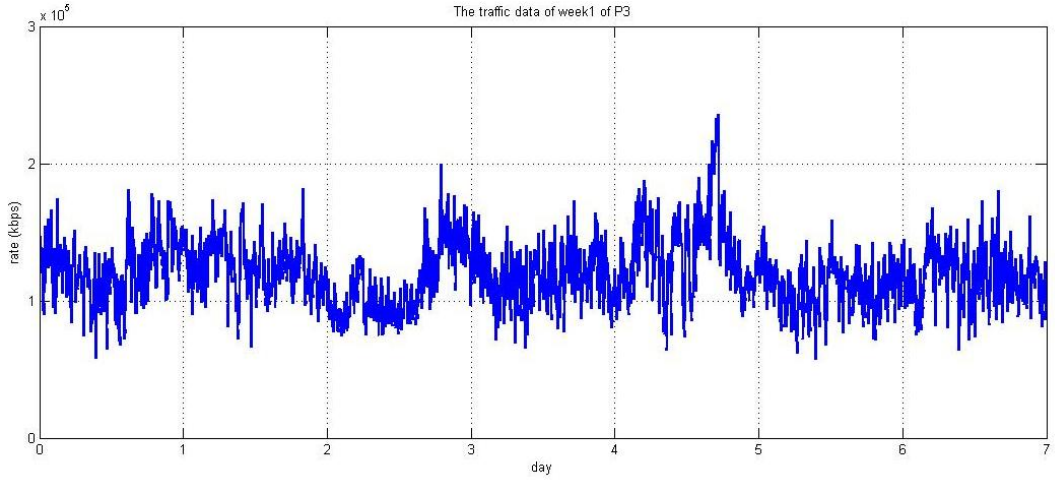

Figure 3.7 First Week Traffic Data of P3

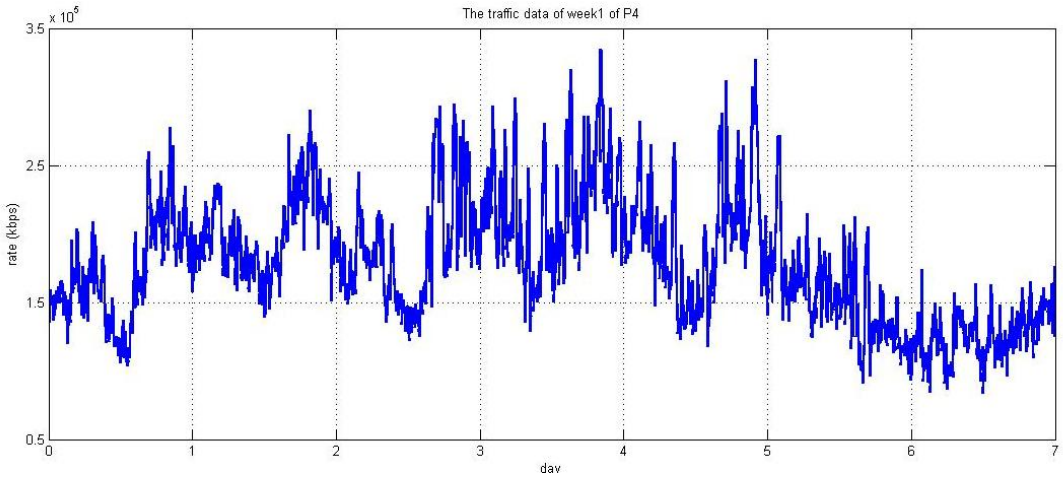

Figure 3.8 First Week Traffic Data of P4

From Figure 3.5 to Figure 3.8, it is clear that different source-destination pairs have different traffic characteristics. For example, P1 has many seemingly random spikes whilst the rest traffic varies little. It seems P1 has a daily pattern. However, compared with the spikes, the remaining P1 traffic has little variation. P2 and P4 both have a noticeable daily pattern with a sinusoid-like waveform shape. In addition, P3 possesses an up / down tendency every day but it is not so significant.

In the first week, the last two days have a lower traffic load than that of the first five days. It is suitable for the weekly pattern characterization that the traffic is high in weekdays and is low at weekends. However, as we are interested in exploring whether the traffic load can be predicted from at roughly the same time (given a certain observation window) over successive days, we removed the data of weekends as they follow a different pattern.

Figure 3.9 - Figure 3.12provide the Sample Auto-Correlation Function (SACF) coefficient of the 4 pairs of weekdays. The data is collected every 5 minutes, so there are 288 data points in a day and 1440 data points over five weekdays. As we want to know if the traffic at roughly the same time has a strong correlation, we used 1440 as the number

of lag and 288 as a step number in the SACF function. The range of SACF coefficient is [-1, 1]. In the range [0, 1], larger values indicate a stronger correlation.

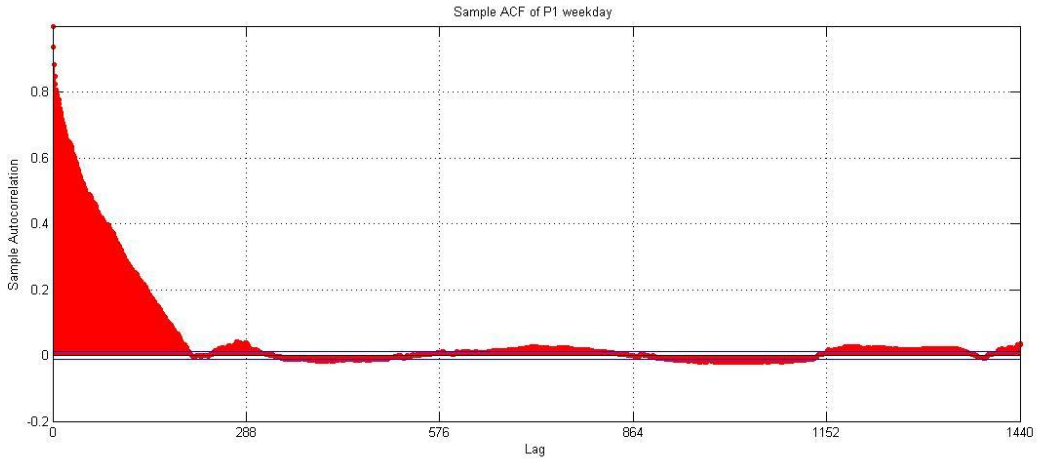

Figure 3.9 Sample ACF of P1

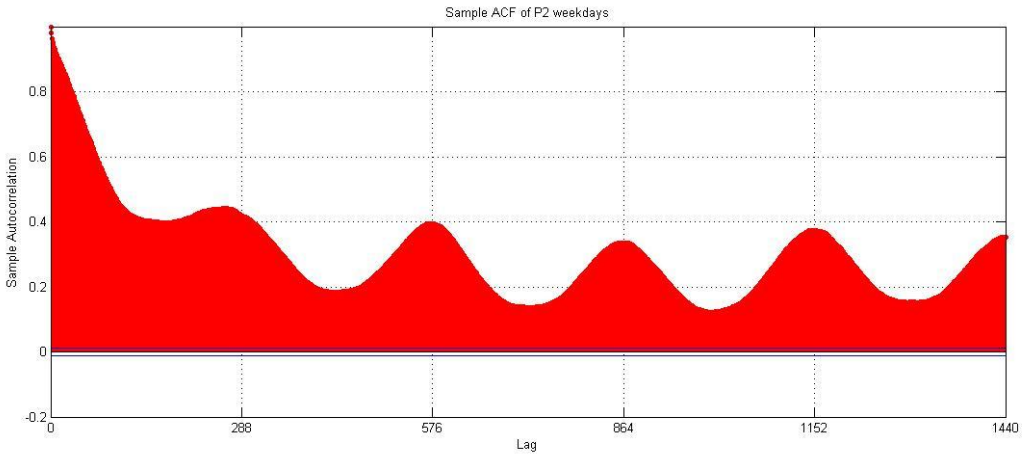

Figure 3.10 Sample ACF of P2

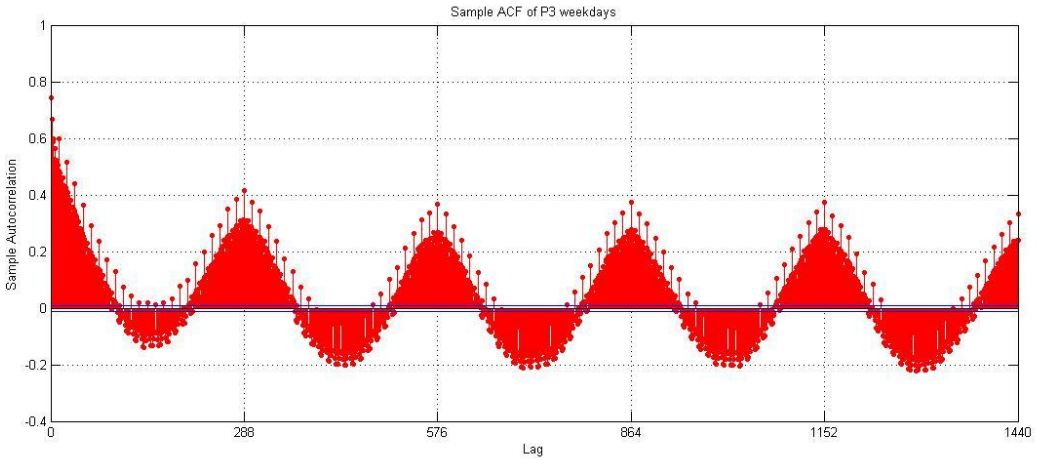

Figure 3.11 Sample ACF of P3

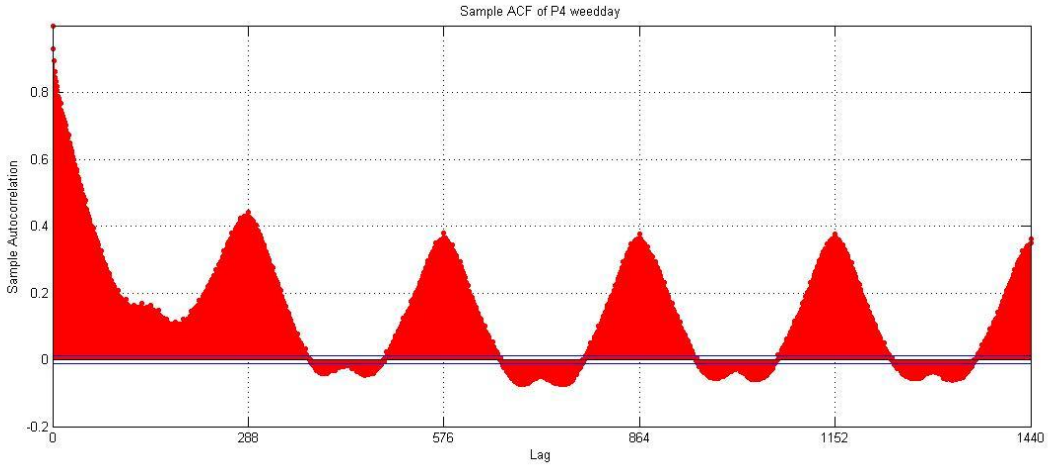

Figure 3.12 Sample ACF of P4

The results of the SACF coefficient of the four pairs are different. P1 has small SACF coefficient when the lag values are 288 (1 day), 576 (2 days), etc. It means that little correlation is shown between the certain traffic and traffic at the same time in the previous day or two days. Thus, we cannot build a P1 model, as a correlation relationship is required in the ARIMA model.

In contrast, P2, P3 and P4 have a high positive correlation coefficient (around 0.4) when lags are equal to 288, 576, 864, 1152 and 1440. This means that for the same time during weekdays, there is a strong correlation. Hence, for a specific time of a day, the amount of traffic has a strong correlation with the values of previous days.

As P2, P3 and P4 have similar characteristic, we selected P2 as an example to perform traffic modelling using MATLAB. As we want to know whether the traffic could be predicted based on the same time on previous days, we used 288 as the difference. Initially, the original data is denoted as x_t and the series is X_t. As there are 119 weekdays and 288 data per day, X_t is represented as:

$$X_t=\{x_1, x_2 \ldots x_{34272}\}$$

The difference between two values at the same time of two days is denoted as y_t, where y_t is represented as

$$y_t = x_t - x_{(t-288)}$$

The series after the difference is denoted as Y_t and Y_t is shown in Figure 3.13.

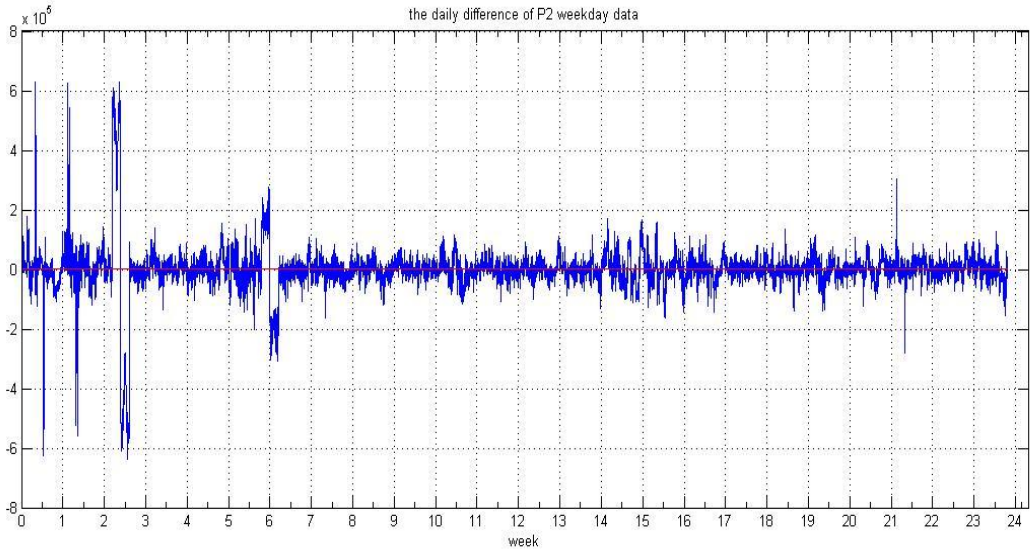

Figure 3.13Daily Difference of P2 Weekday Data

After we differenced the data, we have a series of stationary data. At this stage, we can use several linear models i.e. AR, MA or ARMA models to model the Y_t series and then transfer the prediction of Y_t back to X_t.

At first, different linear models are used to fit Y_t. A metric, called the best fit, is usually used to measure the performance of models. The best fit indicates that the percentage the values of the model that are the same as the raw data. For example, if the best fit is equal to 60, it means that 60% of the values of the model are the same as the raw data. The best fit results of different linear models are shown in Table 3.1. We can see that the best fit values of models are similar. As AR (1) is the simplest model, we selected AR(1) to model the traffic.

Table 3.1TheBest Fit of Models

Model	Best fit
AR(1)	74.13
AR(4)	74.62
AR(10)	74.76
ARMA(1,1)	74.14
ARMA(2,1)	74.55
ARMA(1,2)	74.55

The AR (1) model is presented as follows:

$$y_t = A \times y_{t-1} + e_t$$

where y_t is the value at time t, y_{t-1} is the value of the previous moment. e_t is the white-noise disturbance and A is the coefficient of previous output value. The coefficient of AR (1) is A= 0.9722 . Thus, the AR(1) model of Y series is as follows:

$$y_t = 0.9722 \times y_{t-1} + e_t$$

Then, AR (1) model of Y_t is used to transfer back to X_t.

Since

$$y_t = 0.9722 \times y_{t-1} + e_t$$

$$y_t = x_t - x_{t-288}$$

$$y_{t-1} = x_{t-1} - x_{t-289}$$

So,

$$x_t = x_{t-288} + 0.9722 \times (x_{t-1} - x_{t-289}) + e_t$$

After the model is built, we used half of the raw data as the input and the other half of the data as future traffic data for testing the model. The result shows the best fit is 58.76, which shows that the P2 traffic can be reasonably predicted by the AR (1) model.

Part 3.3 Discussion and Summary

Based on the traffic analysis and modelling in Part 3.1 and Part 3.2, the results show that for a given network, some source-destination pairs have a strong daily pattern and can be reasonably modelled for prediction whilst others display no clear pattern. This may be due to several reasons as follows:

1. The Abilene network is an academic network rather than a commercial network. Thus, it may exhibit different traffic characteristics compared with commercial networks. For example, the traffic demand, the number of hosts and types of service may be different.
2. The Abilene network was upgraded from 2.5 Gbps to 10 Gbps since 2004 and details of the migration strategy are not published. Thus, some influential factors may have impacted on the traffic appearance though this cannot be confirmed.
3. The data of 24 weeks may be not large enough to obtain good estimates; especially, as it is not continuous.

According to the traffic analysis and modelling in Part 3.1 and 3.2, we found that the traffic prediction is more complex than we firstly assumed. In the Abilene network, the traffic between some source-destination pairs can be predicted relatively well whilst this may not be the case in other instances.

Appendix A Reference

[1] H. Leijon, “Basic Forecasting Theories: A Brief Introduction,”, ITU, Tech. Rep., Nov. 1998.

[2] K. Papagiannaki, N. Taft, Z. L. Zhang, and C. Diot, "Long-term forecasting of Internet backbone traffic," Neural Networks, IEEE Transactions on, vol. 16, pp. 1110-1124, 2005.

[3] L. Kleinrock, Queueing Systems, Volume II: Computer Applications, John Wiley & Sons, 1976.

[4] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," ACM SIGCOMM Computer Communication Review, vol. 23, pp. 183-193, 1993.

[5] V. Paxson and S. Floyd, "Wide area traffic: the failure of Poisson modeling," IEEE/ACM Transactions on Networking, vol. 3, pp. 226-244, 1995.

[6] C. Heyde and Y. Yang, "On defining long-range dependence," Selected Works of CC Heyde, pp. 426-431, 2010.

[7] M. E. Crovella and A. Bestavros, "Self-similarity in World Wide Web traffic: evidence and possible causes," Networking, IEEE/ACM Transactions on, vol. 5, pp. 835-846, 1997.

[8] W. Rea, L. Oxley, M. Reale, and J. Brown, "Estimators for long range dependence: an empirical study," Arxiv preprint arXiv:0901.0762, 2009.

[9] T. Karagiannis, M. Molle, and M. Faloutsos, "Long-range dependence ten years of Internet traffic modeling," Internet Computing, IEEE, vol. 8, pp. 57-64, 2004.

[10] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, Time series analysis: Holden-day San Francisco, 1970.

[11] N. K. Groschwitz and G. C. Polyzos, "A time series model of long-term NSFNET backbone traffic," ICC, 1994, pp. 1400-1404 vol. 3.

[12] Y. Shu, Z. Jin, L. Zhang, L. Wang, and O. W. W. Yang, "Traffic prediction using FARIMA models," ICC, 1999, pp. 891-895 vol. 2.

[13] X. Wang and X. Shan, "A wavelet-based method to predict Internet traffic," Communications, Circuits and Systems and West Sino Expositions, 2002, pp. 690-694, vol. 1.

[14] P. Cortez, M. Rio, M. Rocha, and P. Sousa, "Internet traffic forecasting using neural networks," Neural Networks 2006, pp. 2635-2642.

[15] G. Rutka, "Neural Network Models for Internet Traffic Prediction," Electronics and Electrical Engineeging, vol. 4, pp. 55-58, 2006.

[16] “AbileneTM” Available: http://www.cs.utexas.edu/~yzhang/research/AbileneTM/.

[17] “The Abilene Network,”Available: http://www.internet2.edu/pubs/200502-IS-AN.pdf

APPENDIX B: SIMULATION MODELLING

Part 1 Event Routines

1. Src_gen Event

a) Event Description

A *Src_gen* event is used for generating the traffic flow in a source node. In the network, each node (VR instance) is connected with a "source" (Src) and a "destination" (Dest). A Src keeps sending traffic to all Dests. The traffic model is represented as a fluid flow instead of discrete packet level operations. Each traffic flow is labelled with some parameters, e.g. source and destination information. When a traffic flow is generated, the traffic is sent to the location of the connected VR instance. Then, periodically, e.g. at 5 minute intervals, the traffic flow can be updated

A Src_gen event is required to be scheduled initially as the event cannot be scheduled by other events. In the initialization routine, N Src_gen events are created. Then, a Src_gen event can schedule another Src_gen event which will happen after a specific period of time. In addition, a Src_gen event schedules an Arrival event as the traffic is sent to the VR instance. Hence a Src_gen event can reschedule itself as well as an Arrival event.

b) Event Procedure

In the initialization routine, schedule N initial Src_gen events

```
Begin
   For i = 1 to N
      Begin
        Invoke a Src_gen( source name i ) event for Src i
      End
End
```

Function: Src_gen (source name *i*) *//Generate a series of arrival events on Src i for next moment*

```
Begin
   For j = 1 to N
         Begin
           Generate new traffic flow from source i to destination j including related parameters
           Schedule an Arrival event // arrive connected VR instance
      End
      Schedule a Src_gen event(source name i )

  End
```

2. Arrival Event

a) Event Description

An *Arrival* event stands for the operation when the traffic arrives at a VR instance. Then the VR processes and updates the related VR and PP states, e.g. the workload of the VR and the occupied number of line-cards of PP. Next, an *Arrival* event schedules a *Departure* event.

b) Event Procedure

```
Function: Arrival( )

        Begin
            Get parameters from the traffic flow.
            Update the related VR and PP states.
            Schedule a Departure event.
        End
```

3. Departure Event

a) Event Description

After the related VR and PP states are updated in the *Arrival* event, the VR instance forwards the traffic to the next hop. The next hop can be another VR instance or "Dest". If next hop is a VR instance, schedule an *Arrival* event. If it is a Dest, schedule a *Reach_destination* event.

b) Event Procedure

```
Function: Departure( )

        Begin
         Figure out the next hop information
        If(next hop is a VR instance)
                Schedule an Arrival event
        If (next hop is a Dest)
                Schedule an Reach_destination event
        End
```

4. Reach_destination

a) Event Description

When the next hop in a Departure event is a Dest, a Reach_destination event is invoked. The event updates the state of the Dest.

b) Event Procedure

```
Function: Reach_destination( )

        Begin
          Get traffic flow related parameters
          update the state of the destination
        End
```

5. Collect Event

a) Event Description

A *Collect* event is used for collecting the network condition and sending it to the Central Control Unit (CCU). The collected parameters include current workload of PPs, the current workload of VRs, the number of VR line-cards, the state of PPs and the location of VRs. The CCU then examines if the network conditions satisfy the migration conditions. If they do, then schedule a *VRM_MOEA* event. Otherwise, do nothing..

b) Procedure

```
In the initialization routine, schedule a initial Collect events
Function: Collect()
    Begin
        Collect the parameter values from VRs and PPs and store them in CCU
        If  (the network condition satisfies migration conditions )
        Schedule a VRM_MOEA event
    End
```

6. VRM_MOEA

a) Event Description

A*VRM_MOEA* event is used for computing VRM_MOEA algorithm to obtain a group of good solutions for virtual router migration (VRM). A *VRM_MOEA* event always schedules an *Optical_test* event.

b) Procedure

```
    Function: VRM_MOEA()
     Begin
        Compute VRM_MOEA algorithm to obtain a group of good solutions.
        Schedule an OPTICAL_TEST event.
    End
```

7. Optical_test

a) Description

An Optical_test event is used for testing whether the underlying optical layer can support the candidate solutions. If the optical resources can support the candidate solution, a VRM event is scheduled. Otherwise, no VRM will happen

b) Procedure

```
Function: Optical_test ( )
        Begin
   Rank the optimal solutions according to energy saving
            for solution 1 to (size of secondary population)
                Test whether the optical layer supports this candidate solution
                if (optical layer supports)
                    {Schedule a VRM event, End the loop}
          End
```

8. VRM

a) Event description

This event is to move VR instances to their appropriate destination PPs according to the identified candidate solution from the Optical_test event. If some sleeping PPs are needed to be destination PPs and host new VR instance(s), these PPs are re-awoken before the migration. Then, the VR instances are moved to their destination PPs. After the migration, surplus PPs can be put to sleep. Finally, the VR and PP states are updated.

b) Procedure

```
Function: VRM()
    Begin
      Turn on the appropriate PPs if some sleeping PPs are needed to host new VR instance.
      Move VR instances to their destination PPs.
      Turn off the appropriate PPs and release the unneeded optical channels.
    End
```

9. End_simulation

a) Event Description

The event is for terminating the simulation. We use the maximum simulation time as the termination condition. After this event, a report is generated for recording the parameters of interest.

b) Procedure

```
In the initialization routine, schedule a initial End_simulation events
Function: End_simulation( )
    Begin
        Stop the event list simulation
    End
```

Part 2 Samples of Traffic Demand from the Abilene Network

In this section, the traffic load of three weeks of the Abilene network are shown from Figure 2.1 to Figure 2.3.

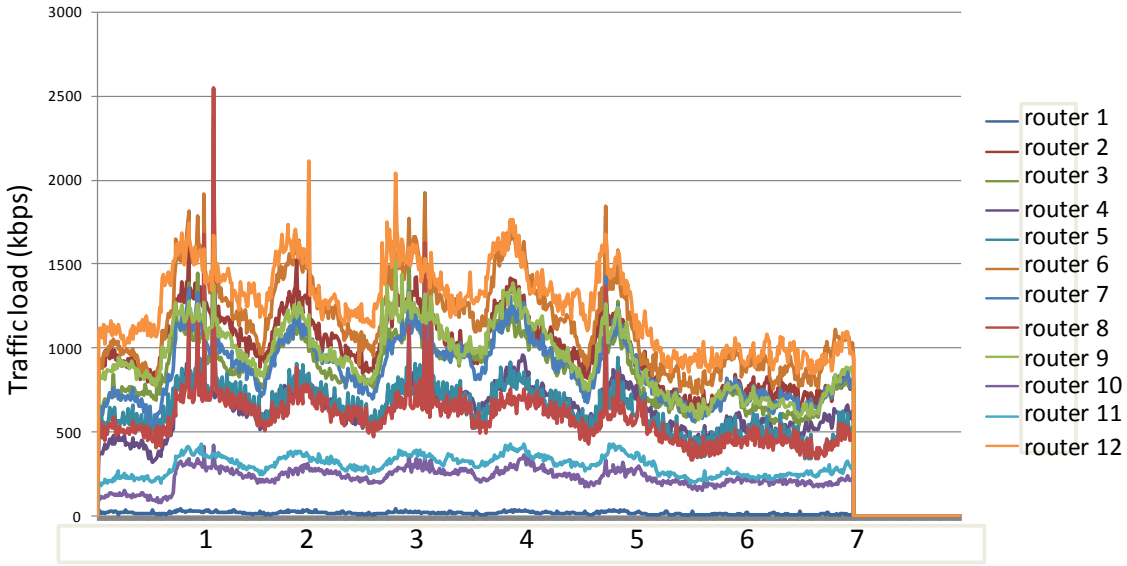

Figure B.2.1 The 1st week traffic load

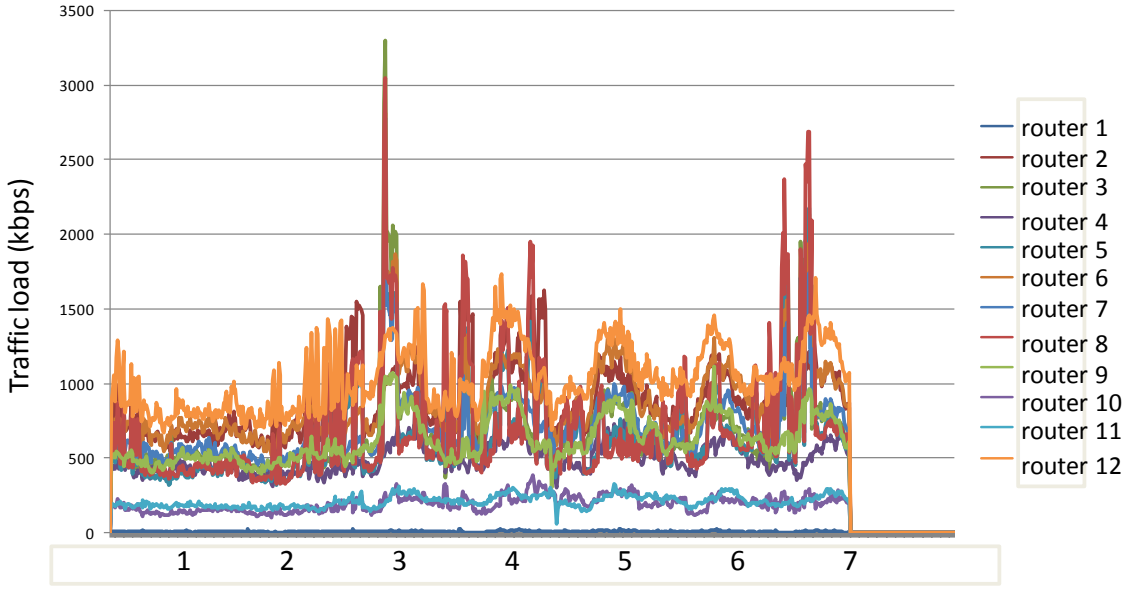

Figure B.2.2. The 13th week traffic load

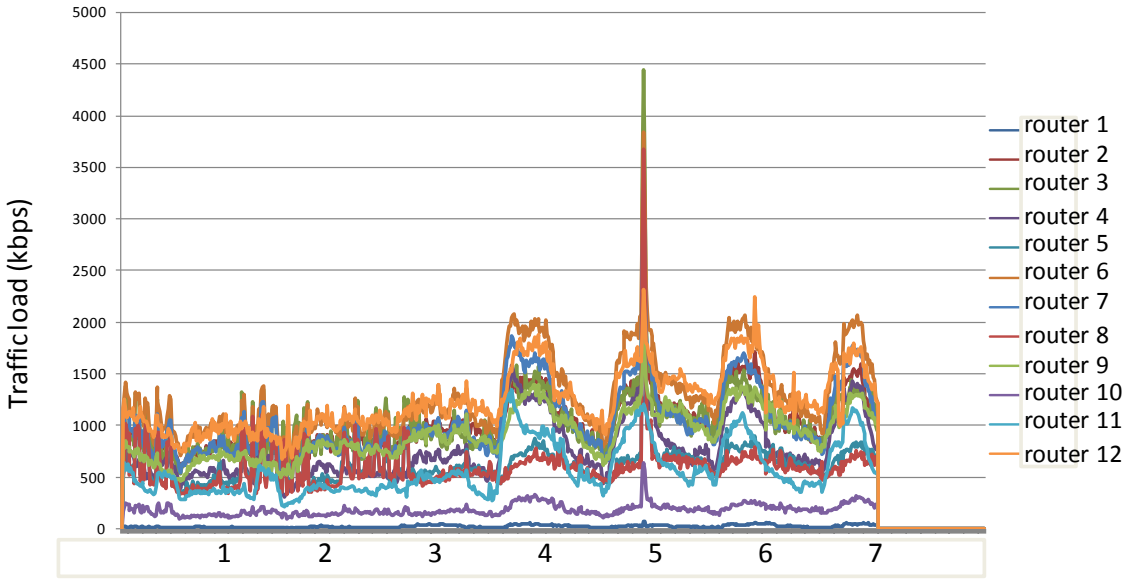

Figure B.2.3. The 24th week traffic load

Part 3 Random Number Generator

The header file

```
/* Header file "mrand.h" to be included by programs using mrand.c */

double mrand(int stream);
void mrandst(double* seed, int stream);
void mrandgt(double* seed, int stream);
```

The c file

```
/* Combined MRG from Sec. 7.3.2, from L'Ecuyer (1999). Multiple(10,000) streams are supported, with
seed vectors spaced 10,000,000,000,000,000 apart. Throughout, input argument "stream"must be an integer
giving the desired stream number. The header filemrand_seeds.h is included here, so must be available in the
appropriate directory. The header file mrand.h must be included inthe calling program (#include "mrand.h")
before using thesefunctions.

   Usage: (Three functions)

   1. To obtain the next U(0,1) random number from stream "stream,"
executeu = mrand(stream);
wheremrand is a double function. The double variable u willcontain the next random number.

   2. To set the seed vector for stream "stream" to a desired 6-vector,execute mrandst(zset, stream);
where mrandst is a void function and zset must be a doublevector with positions 0 through 5 set to the desired
stream 6-vector, as described in Sec. 7.3.2.

   3. To get the current (most recently used) 6-vector of integers inthe sequences (to use, e.g., as the seed for
a subsequentindependent replication), into positions 0 through 5 of thedouble vector zget,
executemrandgt(zget, stream); where mrandgt is void function. */

#include "mrand_seeds.h"
#define norm   2.328306549295728e-10  /* 1.0/(m1+1) */
#define norm2  2.328318825240738e-10  /* 1.0/(m2+1) */
#define m1     4294967087.0
#define m2     4294944443.0

/* Generate the next random number. */
double mrand(int stream)
{
   long k;
   double p,
        s10 = drng[stream][0], s11 = drng[stream][1], s12 = drng[stream][2],
        s20 = drng[stream][3], s21 = drng[stream][4], s22 = drng[stream][5];

   p = 1403580.0 * s11 - 810728.0 * s10;
   k = p / m1;  p -= k*m1;  if (p < 0.0) p += m1;
   s10 = s11;   s11 = s12;  s12 = p;
```

```
    p = 527612.0 * s22 - 1370589.0 * s20;
    k = p / m2;  p -= k*m2;  if (p < 0.0) p += m2;
    s20 = s21;   s21 = s22;  s22 = p;

drng[stream][0] = s10;  drng[stream][1] = s11;  drng[stream][2] = s12;
drng[stream][3] = s20;  drng[stream][4] = s21;  drng[stream][5] = s22;

    if (s12 <= s22) return ((s12 - s22 + m1) * norm);
    else return ((s12 - s22) * norm);
}

/* Set seed vector for stream "stream". */
void mrandst(double* seed, int stream)
{
inti;
    for (i = 0; i<= 5; ++i) drng[stream][i] = seed[i];
}

/* Get seed vector for stream "stream". */
void mrandgt(double* seed, int stream)
{
     inti;
    for (i = 0; i<= 5; ++i) seed[i] = drng[stream][i];
}
```

Part 4 Network Topology Generator Validation

```
The degree of each node :      2    4    4    2    1    3
The connections:
   2    6
   1    3    4    6
   2    4    5    6
   3    2
   3
   1    3    2
```

```
The degree of each node :      4    1    3    3    2    3
The connections:
   2    3    4    6
   1
   1    5    6
   1    5    6
   4    3
   4    1    3
```

```
The degree of each node :      4    3    2    3    3    1
The connections:
   2    3    4    5
   1    4    3
   1    2
```

```
1  5  2
4  6  1
5
```

Part 5 Simulation Tool Validation

In this section, the simulation tool is validated by using a simple 3-node-3-linknetwork and a simple traffic model. The network topology is illustrated in Figure 5.1 and network parameters are shown in Table 5.1.

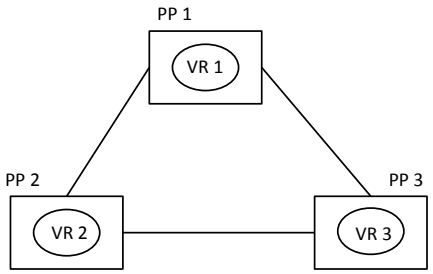

Figure B.5.1 3-node-3-link Network Topology

Table B.5.1 Parameter Values in the 3-node-3-link Network

Parameter name	value	Parameter name	value
Number of PPs	3	Mating pool size	5
Number of VRs	3	Maximum generation	20
Number of ROADMs	3	Crossover rate	0.9
Physical distance between PPs	100 km	Mutation rate	0.1
Number of line-cards	16	VRM time	30 s
Quiet threshold	0.3	Migration cost weight parameter	0.5
Busy threshold	0.7	The fraction of a sleeping PP power consumption	5%
Primary population size	10	Maximum simulation length	1 day

Secondary population size	5	Optical test duration	20s

A simple sinusoidal traffic modelwithout a random noise component is used. The traffic load is sampled every 5 minutes; the maximum traffic load is 30Gbps andγ is equal to 0.2. The traffic equation is shown as follows:

$$\boldsymbol{T_{ij}(t) = 30 \cdot \left[\frac{1-0.2}{2} \cdot \left(1 + \sin\left(\frac{1}{288}t\right)\right) + 0.2\right]}$$

The power consumption values of each device are described in Section 6.3.1. We assume that a PP has 1 LCC and 16 line-cards and a line-card has 40Gbps capacity. Hence, the total switching capacity of a PP is $16 * 40 * 2 = 1280\ Gbps$. The power consumption of devices and the network overall are shown in Table 8.2.

Table B.5.2 Power Consumption of Equipment

Name	Power consumption (W)
PP	16*500 + 1630 = 9630
ROADM	135*2+150 = 420
OLA	110
Overall Network	(9630 + 420+110)*3 = 30408

The network state is collected every 15 minutes. When the simulation clock is equal to 15 minutes from the start, three PPs are in quiet mode as the traffic load is light. Hence, the VRM_MOEA is computed to select the appropriate destination physical platform(s) for saving energy. The result of VRM_MOEA is shown in Figure 8.2. As we can see in Figure 8.2, there are three non-dominated solutions in the final secondary population based on their fitness values. Specifically, the candidate solution “1 2 3” is the same as current network configuration. As the migration cost of the solution “1 2 3” is 0 which is the smallest in the population, it survives. The power consumption and migration cost of the remaining two solutions are discussed as follows. Firstly, because all ROADMs and OLAs remain working during the simulation, we only consider the power consumption of PPs in the evaluation stage of VRM_MOEA. When three VRs are running on the same PP in the candidate solution “1 1 1” and PP 2 and PP 3 can be put into the sleep state for saving energy, the power consumption is $9630 + (1630 * 0.05 * 2) = 9793\ W$. Similarly, power consumption of two working PPs and one sleeping PP in the candidate solution “1 2 1” is$9630 * 2 + (1630 * 0.05) = 19341.5\ W$.

Then, the VRM cost is composed of two components: the hop-distance from the default PP location to the destination PP location and current PP location to the

destination PP location. The migration cost of candidate solution "1 1 1" is calculated based on Equation 4.5 as follows:

$$Cost = \mathrm{w_{cost}} \cdot Cost_1(I) + (1 - \mathrm{w_{cost}}) \cdot Cost_2(I) = 0.5 * 2 + 0.5 * 2 = 2$$

Similarly, the candidate solution "1 2 1" has a migration cost 1.

start VRM_MOEA at 900.5 second

Secondary population at Generation = 20:

The generation,20

8, 1, 1, 1, 9793.000, 2.000, 0.000

6, 1, 1, 1, 9793.000, 2.000, 0.000

5, 1, 2, 1, 19341.500, 1.000, 0.000

6, 1, 2, 3, 28890.000, 0.000, 0.000

9, 1, 2, 3, 28890.000, 0.000, 0.000

Rank solutions based on power consumption:

The generation,20

8, 1, 1, 1, 9793.000, 2.000, 0.000

6, 1, 1, 1, 9793.000, 2.000, 0.000

5, 1, 2, 1, 19341.500, 1.000, 0.000

6, 1, 2, 3, 28890.000, 0.000, 0.000

9, 1, 2, 3, 28890.000, 0.000, 0.000

Eliminate duplicate solutions:

The generation,20

8, 1, 1, 1, 9793.000, 2.000, 0.000

6, 1, 2, 1, 19341.500, 1.000, 0.000

5, 1, 2, 3, 28890.000, 0.000, 0.000

6, -1, -1, -1, -1.000, -1.000, -1.000

9, -1, -1, -1, -1.000, -1.000, -1.000

The computational time of VRM_MOEA is 0.04 seconds

Figure B.5.2 VRM_MOEA Result

After VRM_MOEA is complete, candidate solutions are sent to the optical resource test unit. The solutions are tested one by one based on their power consumption rank. The first one which passes the test will be executed as the appropriate destination arrangement for virtual router migration. In this case, candidate solution "1 1 1" is supported by the optical resource and VR 2 and VR 3 are moved to PP1.After VRM, 3 VRs are running on the same PP. In this case, because the VR traffic load is light throughout the day, 3 VRs remain on PP1 for the remaining time.

The baseline scenario energy consumption of a day is thus:

$$30408 * 24 * 60 * 60 = 2627251200\ J$$

The energy consumption of energy-efficient scenario is:

$$30408 * (900.5 + 20 + 30) + (9793 + 420 * 3 + 110 * 3) * (24 * 60 * 60 - (900.5 + 20 + 30)) = 1001574463\ \text{J}$$

The energy saving percentage is:

$$(2627251200 - 1001574463)/2627251200 = \boldsymbol{61.8\%}$$

The simulation result is the same as the analytical result, indicating that the simulation model works as expected.

APPENDIX C: SIMULATION RESULTS

Part 1 BLX-a Results

As one single point crossover has little impact on the performance of VRM_MOEA, we repeat the simulations with a popular crossover mechanism, BLX-a. The results are shown from Figure c.1.1 to Figure c.1.3. The results have a similar trend to the corresponding figure of one point crossover mechanism.

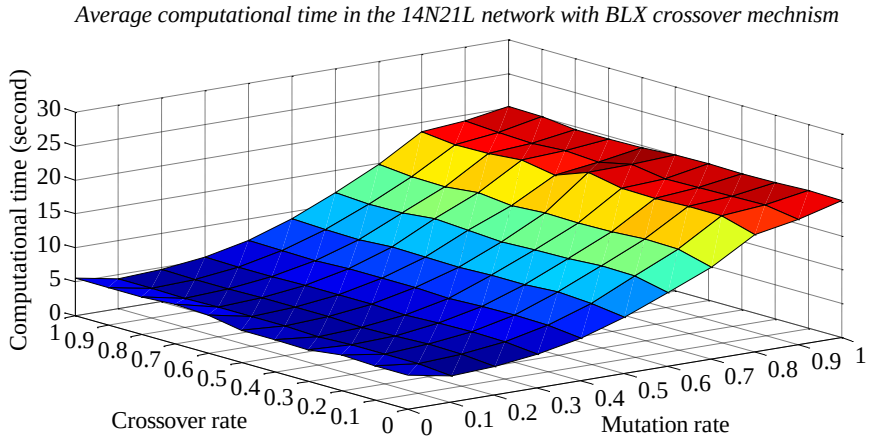

Figure c.1.1 Computational Time of Various Mutation and Crossover Rate

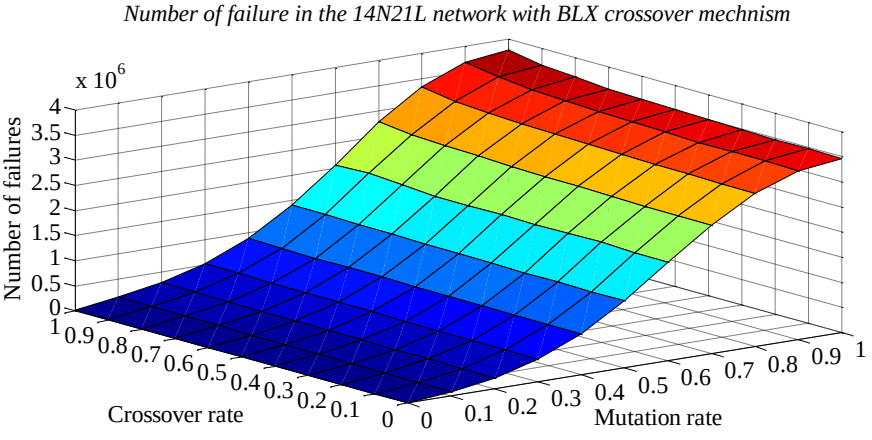

Figure c.1.2 Accumulated Number of Failures of Viability Test

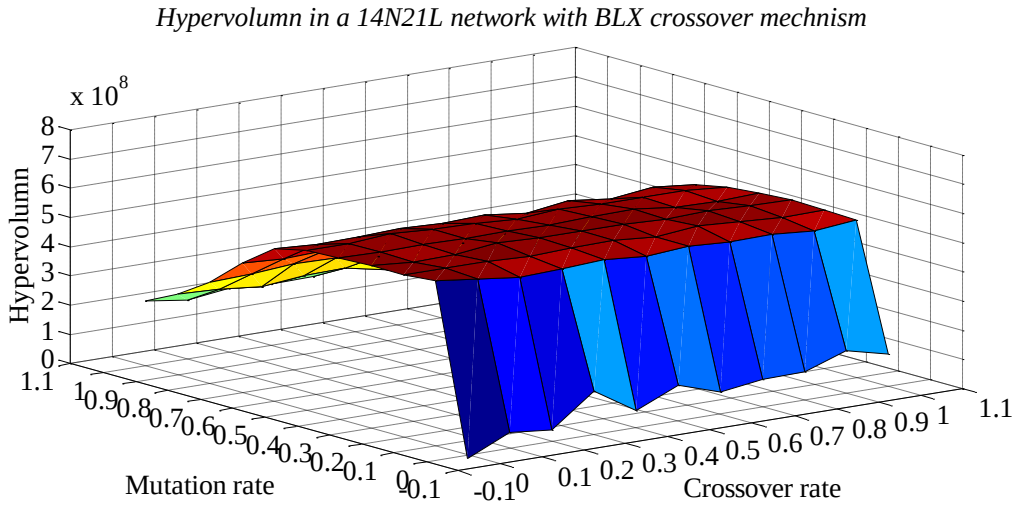

Figure c.1.3 *Hypervolume* with Various Mutation and Crossover Rate

Part 2 Migration Cost Setting

Figure c.2.1 and Figure c.2.2 shows the number of occupied optical channels in the Abilene network in the 13th week and 24th week with different cost settings, respectively.

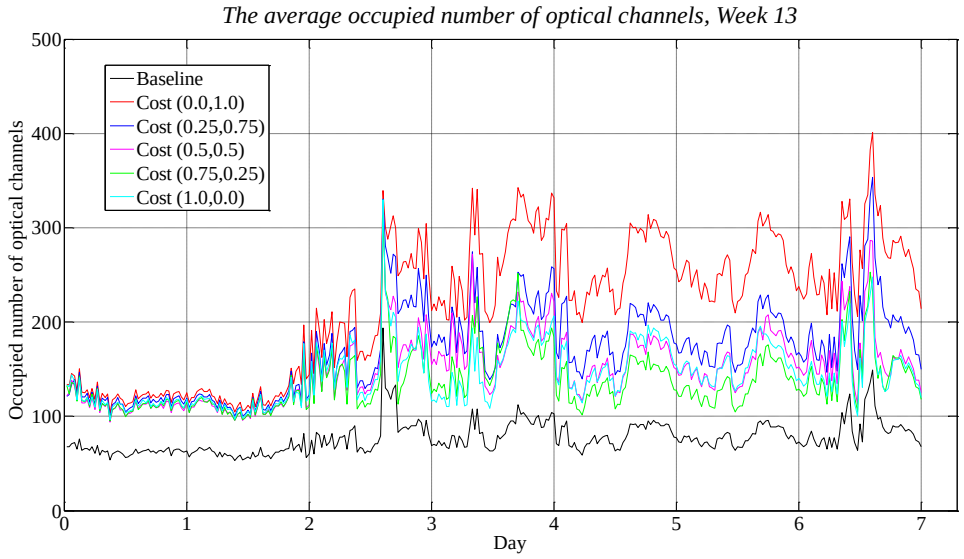

Figure c.2.1 The Number of Occupied Optical Channels in the Abilene Network in the 13rd Week

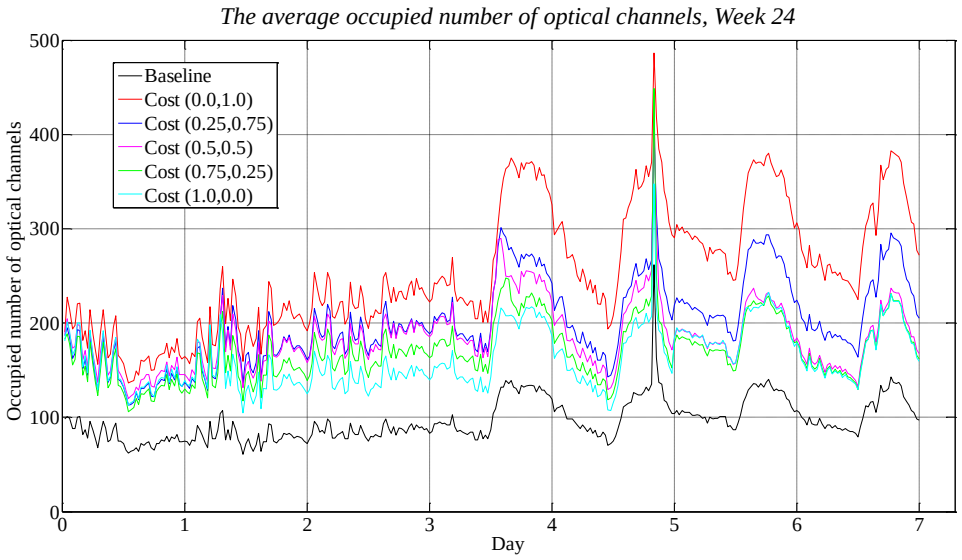

Figure c.2.2 The Number of Occupied Optical Channels in the Abilene Network in the 24th Week