

# Policy-driven Traffic Engineering in Energy-aware ISP Backbone Networks

Frederic Francois

Submitted for the Degree of  
Doctor of Philosophy  
from the  
University of Surrey



Centre for Communication Systems Research  
Faculty of Electronics and Physical Sciences  
University of Surrey  
Guildford, Surrey GU2 7XH, UK

October 2013

© Frederic Francois 2013

# Summary

The excessive energy consumption of backbone networks is causing concerns among network operators. This thesis focuses on the design of Energy-aware Traffic Engineering (ETE) schemes which improve the energy-efficiency of different backbone networks by enabling the delivery of traffic by the smallest number of network devices so that the remaining devices can go to sleep during the periods of low traffic demands.

The first proposed ETE scheme is called *Time-driven Link Sleeping* (TLS) which uses only two network routing topologies: the full topology with all links being active, and a reduced one with a subset of links sleeping. The key novelty of TLS lies in its ability to jointly optimize the reduced network topology and the off-peak period during which it is operated. Moreover, an extension to TLS makes it robust to single link failures. The second ETE scheme is a *Green Load-balancing Algorithm* (GLA) which complements TLS and other existing ETE schemes by jointly optimizing the IGP link weights in backbone networks for improved load-balancing and energy-efficiency after these existing ETE schemes put links to sleep.

The final contribution is an online distributed ETE scheme called *Green Backup Paths* (GBP) which dynamically diverts traffic from some selected links onto their backup paths, which were pre-installed to protect against link failure, so that these links have the opportunity to go to sleep without affecting the primary purpose of the backup paths. The distributed nature of GBP makes it scalable to large networks and be very responsive to sudden traffic changes since multiple routers can concurrently make interference-free decisions.

The simple TLS scheme with GLA is ideally suited for networks which experience a regular traffic pattern because of their offline nature while the more complex GBP scheme is more suitable when there is dynamic traffic because of its online nature.

**Keywords:** Energy-efficiency, Load-balancing, Policy, Maximum Packet Delay, Traffic Engineering, Offline Optimization, Distributed Network Optimization, Online Optimization and Quality-of-Service.

Email: f.francois@surrey.ac.uk

WWW: <http://personal.ee.surrey.ac.uk/Personal/F.Francois/>

# Acknowledgments

I would like to take this opportunity to express my sincere gratitude to my PhD supervisor Dr. Ning Wang. He has given extremely valuable advice and unfaltering support throughout the three years of this PhD. Our valuable collaboration has allowed me to complete this PhD on time as well as write three journal papers and three conference papers, including one best student paper award. I am also grateful for the patience he has shown to me during the PhD.

In addition, I would like to thank Dr. Stylianos Georgoulas for his valuable help and insightful discussions during the PhD. He has spent a generous amount of his time discussing the various technical aspects of the different technical works in this PhD thesis.

I would like to thank Prof. Klaus Moessner who first recruited me to study at the University of Surrey. He has generously provided funding without which this PhD would not be possible. I am grateful that he took time to read this thesis and the valuable advice that he gave.

I would like to thank my medical team, in particular Dr. Hopkirck, Ms. Hall, Mr. Fenwick-Smith and Ms. Peacock, who have helped me greatly when I was sick. They have always been very supportive and patient in their search of an appropriate solution for me.

I would like to thank my family for the support and kindness that they have shown to me throughout my life. I wish to particularly express my gratitude to my mum for all these years of emotional as well as financial support.

I would like to thank my friends for all the support that they have generously provided during my time in Guildford: Bambang, Chamitha, Daphne, Dinesha, Donann, Glenn, Imalka, Kayode, Mohammed, Osman, Ranga, Ricardo, Shoban and Yashil. I would like to also thank the administrative and technical support team at University of Surrey.

Last but not least, I wish to thank Dr. Michael Howarth and Prof. David Hutchison for acting as my PhD exam committee. They have given valuable suggestions to improve this PhD thesis.

# Contents

<b>Summary.....</b>	<b>ii</b>
<b>Acknowledgments .....</b>	<b>iii</b>
<b>Contents .....</b>	<b>iv</b>
<b>List of Figures.....</b>	<b>viii</b>
<b>List of Tables .....</b>	<b>x</b>
<b>Glossary of Terms.....</b>	<b>xii</b>
<b>List of Notations .....</b>	<b>xv</b>
<b>Thesis Related Publications .....</b>	<b>1</b>
<b>Journal Publications .....</b>	<b>1</b>
<b>Conference Publications .....</b>	<b>1</b>
<b>1 Introduction.....</b>	<b>2</b>
1.1 Background and Motivation.....	2
1.2 Novel Energy-aware Traffic Engineering Schemes.....	3
1.3 Thesis Structure .....	7
<b>2 Literature Review .....</b>	<b>9</b>
2.1 Introduction.....	9
2.2 Energy Consumption of Network Devices and Components .....	10
2.3 Network Energy Saving Strategies .....	11
2.4 Energy-aware Traffic Engineering (ETE).....	12
2.4.1 ETE Classification.....	12
2.4.2 ETE Analysis.....	17
2.4.2.1 Offline ETE Schemes Analysis.....	17
2.4.2.2 Online ETE Schemes Analysis .....	24
2.4.2.3 Sleeping versus Rate Adaptation .....	30
2.5 Conclusions.....	31
<b>3 Optimization of Time-driven Link Sleeping Reconfigurations in ISP Networks ...</b>	<b>34</b>
3.1 Introduction.....	34

3.2	Time-driven Link Sleeping (TLS) Overview.....	35
3.3	Problem Formulation .....	37
3.4	Proposed Heuristic Scheme .....	38
3.4.1	Stage 1: Computation of the synthetic TM and the starting point for off-peak window size expansion .....	39
3.4.2	Stage 2: Greedy link removal .....	40
3.4.3	Stage 3: Determination of off-peak window size and the final reduced topology .....	42
3.5	Performance Evaluation.....	43
3.5.1	Experiment Setup .....	43
3.5.2	Simulation Results.....	45
3.6	TLS with Single Link Failure Protection (TLS-SLFP).....	52
3.7	Problem Formulation for TLS-SLFP .....	53
3.8	Modifications to the basic TLS Scheme for TLS-SLFP .....	53
3.8.1	Modified Stage 2: Greedy link removal with full network connectivity check for single link failure scenarios.....	53
3.8.2	New Stage 4: Optimizing sleeping link selection to avoid traffic congestion during single link failures .....	54
3.9	Performance Evaluation.....	56
3.10	Implementation by Network Operators.....	59
3.11	Summary .....	59
<b>4</b>	<b>Optimizing IGP Link Weights for Energy-efficiency and Load-balancing in ISP Networks .....</b>	<b>61</b>

4.1	Introduction.....	61
4.2	Problem Formulation and Existing ETE Schemes.....	64
4.2.1	Problem Formulation.....	64
4.2.2	Existing ETE Schemes .....	65
4.2.2.1	The Least Flow Scheme.....	65
4.2.2.2	The Most Power Scheme .....	65
4.2.2.3	The Time-driven Link Sleeping Scheme .....	66
4.3	Green Load-balancing Algorithm (GLA) .....	67
4.3.1	Scheme Overview.....	67
4.3.2	Solution Encoding .....	68
4.3.3	Fitness Functions.....	68
4.3.4	Sleeping Link Crossover Operator .....	68
4.3.5	Link Utilization Mutation Operator.....	69
4.3.6	Overall Operation of GLA.....	70

4.4 Performance Evaluation .....	70
4.4.1 Network Scenario .....	70
4.4.2 Simulation Results.....	71
4.5 Enhanced GLA for TLS.....	75
4.5.1 Solution-enhancement Heuristic .....	75
4.5.2 Evaluation.....	77
4.6 Maximum Packet Delay (MPD) Performance of the existing ETE Schemes.....	77
4.7 Improvement to current ETE Schemes for Maximum Packet Delay Constraint .....	80
4.8 Performance Evaluation of ETE Schemes with Maximum Packet Delay Constraint.....	81
4.8.1 Performance Comparison between the Original and MPD-aware ETE Schemes .....	81
4.8.2 Performance Comparison between Default and GLA link weights for MPD-aware ETE Schemes .....	82
4.9 Summary .....	84

## **5 Leveraging MPLS Backup Paths for Distributed Energy-aware Traffic Engineering.....**.....**86**

5.1 Introduction.....	86
5.2 Problem Formulation .....	89
5.3 Green Backup Paths (GBP).....	90
5.3.1 Scheme Overview.....	90
5.3.2 Offline Component.....	93
5.3.2.1 Identification of Eligible Backup Paths .....	93
5.3.2.2 Generation of Interference-Risk Links Lists.....	93
5.3.3 Online Component.....	94
5.3.3.1 Stage 1: Gathering the State of the Network.....	95
5.3.3.2 Stage 2: Deactivation of Overloaded Backup Paths .....	96
5.3.3.3 Stage 3: Selection of Multiple Token Holding Links .....	96
5.3.3.4 Stage 4: Offloading of the Token Holding Link .....	99
5.3.4 Handling of Logical Link Failures when GBP is active.....	103
5.4 GBP Performance Evaluation .....	106
5.4.1 Network Scenarios.....	106
5.4.2 Power and Energy Saving Gains .....	106
5.4.3 Maximum Logical Link Utilization.....	109
5.4.4 Increase in Maximum Packet Delay.....	110
5.4.5 Single Logical Link Failure Analysis.....	112
5.4.5.1 Post-failure Peak Maximum Logical Link Utilization.....	112
5.4.5.2 Impact on Energy Saving Gains during Single Logical Link Failures .....	114

5.5 Summary .....	116
<b>6 Conclusions .....</b>	<b>117</b>
6.1 Thesis Summary.....	117
6.2 Thesis Conclusions .....	120
6.3 Future Work .....	121
<b>Bibliography .....</b>	<b>122</b>

# List of Figures

Figure 1-1: Predictable and unpredictable traffic patterns. ....	4
Figure 2-1: Selection criteria for router and link sleeping in [17], [24], [26], [29] and [46] .....	18
Figure 2-2: Illustrative ISP network with backbone, aggregation and edge segments.....	19
Figure 2-3: Network where a bundle of physical links connects any two routers.....	20
Figure 2-4: Network where unnecessary optical-to-electrical domain conversions are avoided. ....	22
Figure 2-5: The different operating modes of a router proposed in [21].....	22
Figure 2-6: Energy profiles of current and future network devices.....	24
Figure 2-7: The two operating states of Power Saving Routers.....	25
Figure 2-8: The timeline for reserving a timeslot and bandwidth for a bulk transfer of data.....	26
Figure 2-9: Step power profile. ....	29
Figure 2-10: Internal function blocks of an AWM gateway. ....	29
Figure 3-1: Illustration of the overall approach. ....	39
Figure 3-2: Flow chart for greedy link removal.....	41
Figure 3-3: Flow chart showing the “off-peak window determination” stage. ....	42
Figure 3-4: Maximum Link Utilization versus Traffic Matrix Index for GÉANT. ....	44
Figure 3-5: Variation of MLU with reduced topology of $\alpha=90$ .....	48
Figure 3-6: Variation of MLU with reduced topology of $\alpha=80$ .....	49
Figure 3-7: Variation of MLU with reduced topology of $\alpha=70$ .....	49
Figure 3-8: Variation of MLU with reduced topology of $\alpha=60$ .....	50
Figure 3-9: Illustrative network topology to demonstrate the reduction of the MLLU with the reduced topology. ....	50
Figure 3-10: Percentage utilization of links in the network.....	51
Figure 3-11: Flow chart showing optimization of the set of sleeping links in TLS-SLFP.....	55
Figure 3-12: Energy saving gains with reduced topology of $\alpha=90$ under different traffic scaling factors. ....	56
Figure 3-13. Energy saving gains with reduced topology of $\alpha=80$ under different traffic scaling factors. ....	57
Figure 3-14. Energy saving gains with reduced topology of $\alpha=70$ under different traffic scaling factors. ....	57
Figure 3-15. Energy saving gains with reduced topology of $\alpha=60$ under different traffic scaling factors. ....	58

---

Figure 4-1: Illustrative network topology to illustrate optimization of link weights.....	63
Figure 4-2: Sleeping link crossover operator.....	69
Figure 4-3: Energy-efficiency of TLS using the different sets of link weights.....	73
Figure 4-4: The MLU variation across 5 days for Default, IGP-WO and GLA link weights when $\alpha$ is set equal to $U$ .....	74
Figure 4-5: The MLU variation across 5 days for GLA link weights when $\alpha$ is set equal to 60, which is below $U$ .....	74
Figure 4-6: Flow chart describing the operation of the Solution-enhancement Heuristic.....	75
Figure 4-7: Comparison of energy-efficiency between GLA and GLA-SH.....	77
Figure 5-1: Basic network topology to illustrate how links are protected in MPLS backbone networks.....	88
Figure 5-2: Timeline for the online operation of GBP.....	92
Figure 5-3: Basic network topology to illustrate how IRLLs are generated.....	94
Figure 5-4: Flow chart showing the different operations in one GBP optimization cycle.....	95
Figure 5-5: Flow chart showing the deactivation of overloaded backup paths.....	97
Figure 5-6: Flow chart showing the generation of the two sub-lists: <i>priority_links_list</i> and <i>normal_links_list</i> .....	98
Figure 5-7: Flow chart showing how to select multiple TH links.....	99
Figure 5-8: Flow chart showing the overall process of offloading a TH link.....	101
Figure 5-9: Flow chart showing the offloading of <i>b_flows</i> from a TH link.....	102
Figure 5-10: Flow chart showing the offloading of <i>p_flows</i> from a TH link.....	103
Figure 5-11: Illustrative topology to demonstrate the conventional failure-protection mechanism in an MPLS-enabled network.....	104
Figure 5-12: Illustrative topology to demonstrate GBP enhanced failure-protection mechanism.....	105
Figure 5-13: Power saved for the GÉANT topology.....	108
Figure 5-14: Power saved for the Abilene topology.....	108
Figure 5-15: Variation of MLLU for Original and GBP for GÉANT topology.....	111
Figure 5-16: Variation of MLLU for Original and GBP for Abilene topology.....	111
Figure 5-17: Post-failure peak MLLU between no-GBP and GBP operations for the GÉANT topology.....	113
Figure 5-18: Post-failure peak MLLU between no-GBP and GBP operations for the Abilene topology.....	113

# List of Tables

Table 2-1: Breakdown of the energy consumption of a typical backbone router.....	13
Table 2-2: Taxonomy of existing offline ETE schemes. ....	14
Table 2-3: Taxonomy of existing online ETE schemes. ....	16
Table 3-1: Definition of symbols for TLS. ....	37
Table 3-2: Link characteristic of GÉANT network topology. ....	44
Table 3-3: Characteristics of 6 weeks of the GÉANT traffic matrices. ....	44
Table 3-4: Energy-efficiency of proposed algorithm.....	47
Table 3-5: Increase in maximum packet delay due to reduced topology.....	47
Table 3-6: Traffic demands and paths for the illustrative network topology in Figure 3-9.....	50
Table 3-7: Performance of week 0 configuration on other weeks. ....	52
Table 3-8: Energy-efficiency of TLS-SLFP.....	58
Table 4-1: Definition of symbols for GLA. ....	64
Table 4-2: Power used due to an active link. ....	71
Table 4-3: Performance comparison of three sets of link weights for LF ETE scheme.....	72
Table 4-4: Performance comparison of three sets of link weights for MP ETE scheme.....	73
Table 4-5: Performance of GLA for the three set of link weights. ....	73
Table 4-6: Performance comparison between GLA and GLA-SH. ....	78
Table 4-7: Increase in maximum packet delay for the reduced network topology due to the Default, IGP-WO and GLA link weights for LF ETE scheme. ....	79
Table 4-8: Increase in maximum packet delay for the reduced network topology due to the Default, IGP-WO and GLA link weights for MP ETE scheme. ....	79
Table 4-9: Increase in maximum packet delay due to the Default, IGP-WO and GLA-SH link weights for TLS ETE scheme.....	80
Table 4-10: Increase in MPD for the full network topology due to the IGP-WO link weights compared to the Default ones. ....	82
Table 4-11: Change in energy-efficiency when LF-D and MP-D, rather than LF and MP, are applied to the Default link weights.....	82
Table 4-12: Change in the improvement of the load-balancing and energy-efficiency when LF-D and MP-D, rather than LF and MP, are applied to the GLA link weights. ....	82
Table 4-13: Performance improvement of the load-balancing and energy-efficiency when TLS-D, rather than TLS, is applied to the GLA-SH link weights.....	83

Table 4-14: Performance improvement when GLA link weights are used for the LF-D and MP-D ETE scheme compared to Default link weights.....	83
Table 4-15: Performance improvement when GLA-SH link weights are used for the TLS-D ETE scheme compared to Default link weights. ....	83
Table 5-1: Definition of symbols for GBP.....	89
Table 5-2: Description of all notions used in GBP.....	91
Table 5-3: Power model of GÉANT network topology.....	107
Table 5-4: Power model of Abilene network topology.....	107
Table 5-5: Energy-efficiency, $E$ , for the GÉANT and Abilene topology.....	110
Table 5-6: Increase in maximum packet delay for the GÉANT and Abilene topology.....	112
Table 5-7: Change in energy-efficiency, $\Delta E$ , for the GÉANT and Abilene topology. ....	115

# Glossary of Terms

AS	Autonomous System
AWM	Active Window Management
BGH	Bi-level Greedy Heuristic
CDN	Content Centric Network
DFS	Dynamic Frequency Scaling
DLFR	Distributed Loop-Free Routing
DPRA	Dijkstra-based Power-aware Routing Algorithm
DVS	Dynamic Voltage Scaling
EARTH	Energy Aware service Rate Tuning Handling
EGH	Exhaustive Greedy Heuristic
EPRA	Energy Profile-aware Routing
ER	Erdos and Renyi
ETE	Energy-aware Traffic Engineering
FGH	Fast Greedy Heuristics
GBP	Green Backup Paths
GDRP-PS	General Distributed Routing Protocol for Power Saving
GLA	Green Load-balancing Algorithm
GLA-SH	Green Load-balancing Algorithm with Solution-enhancement Heuristic
HS	Harmonic Series
ICN	Information Centric Network
ID	Identifier
IGP	Interior Gateway Protocol
IGP-WO	Interior Gateway Protocol Weight Optimizer
IP	Internet Protocol
IRLL	Interference-Risk Links List
IS-IS	Intermediate System-Intermediate System
ISP	Internet Service Provider
LF	Least-Flow
LF-D	Least Flow scheme with Maximum Packet Delay constraint
LK	Least-link
LLN	Least Loaded Node
LLN-WS	Least Loaded Node with Weight Setting
LNF	Lightest Node First

LNF	Lightest Node First
LNF-WS	Lightest Node First with Weight Setting
LR	Lagrangian Relaxation
LSA	Link State Advertisement
LSP	Label Switched Path
MILP	Mixed Integer Linear Programming
MILP-EWO	Mixed Integer Linear Programming with Energy-aware Weight Optimization
MLLU	Maximum Logical Link Utilization
MLU	Maximum Link Utilization
MP	Most Power
MPD	Maximum Packet Delay
MP-D	Most Power scheme with Maximum Packet Delay constraint
MPLS	Multi-Protocol Label Switching
MST	Minimum Spanning Tree
MT	Multi-Topology
MT-ID	Multi-Topology Identifier
NP	Non-deterministic Polynomial
NSGA	Non-dominant Sorting Genetic Algorithm
OE	Opt-Edge
OPEX	OPerational Expenditure
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
PL	Power Law
PoP	Point-of-Presence
PSR	Power Saving Router
QoS	Quality of Service
R	Random
ROD	Routing-On-Demand
SD	Source-Destination
SH	Solution-enhancement Heuristic
SLFP	Single Link Failure Protection
TCP	Transmission Control Protocol
TE	Traffic Engineering
TE-LSA	Traffic Engineering Link State Advertisement
TH	Token Holding
TLS	Time-driven Link Sleeping

TLS-D	Time-driven Link Sleeping scheme with Maximum Packet Delay constraint
TLS-SLFP	Time-driven Link Sleeping with Single Link Failure Protection
TM	Traffic Matrix
TUB	Theoretical Upper Bound
WDM	Wavelength Division Multiplexing
WS	Watts-Strogatz

# List of Notations

$A_x$	Set of sleeping links in the reduced topology during off-peak time
$\alpha$	Maximum allowable utilization of link capacity
$\beta$	Scaling factor for traffic matrix
$c_{ij}$	Capacity of link from router $i$ to $j$
$D$	Maximum packet delay
$E$	Energy-efficiency
$\varepsilon$	Fractional multiplier less than 1
$f_{ij}$	Total traffic demand on link from router $i$ to $j$
$f_{ij}^{sd}$	Traffic demand from router $s$ to $d$ that traverses link from router $i$ to $j$
$G(R, L)$	Directed graph with $R$ being set of routers and $L$ being set of links
$H$	Number of off-peak traffic matrices which produce an $U$ above $\alpha$
$h_l$	Spare capacity of logical link $l$
$I_x$	Set of consecutive traffic matrices supported by reduced topology $A_x$
$K$	Number of paths for each Source-Destination paths in the network
$\kappa$	Number of objectives
$L$	Set of (logical) links
$M_{v,q}$	$q^{\text{th}}$ traffic matrix of $v^{\text{th}}$ day
$N$	Normal distribution
$\Omega$	Maximum allowable increase in maximum packet delay
$\Phi_l$	Excess load on logical link $l$
$\Psi$	Maximum amount by which the load-balancing value given by GLA results can be above the minimum load-balancing value given by GLA.
$\Pi(t^{sd})$	Set of links taken by $t^{sd}$ from router $s$ to $d$
$P$	Power consumed
$p_l$	Power consumed by one physical link in the logical link $l$
$Pr_{mut}$	Probability of mutation of selected chromosome
$Pr_{g\_mut}$	Probability that a gene in a mutating chromosome will mutate
$q$	Traffic matrix index in a day
$R$	Set of routers
$t^{sd}$	Traffic demand from router $s$ to $d$
$T$	Total time

$\theta$	Number of traffic matrices from the expansion point
$T_o$	Off-peak time duration
$U$	Maximum link utilization
$v$	Day index, with 0 being Monday, 1 being Tuesday, etc.
$W_n$	New link weight
$W_o$	Original link weight
$\chi$	Population size of Green Load-balancing Algorithm
$x$	Indexer
$\lambda$	Optical-carrier transmission speed
$y_l$	Total number of physical links in logical link $l$
$y_l^o$	Number of sleeping physical links in logical link $l$
$z_l$	Capacity of one physical link in logical link $l$

# Thesis Related Publications

## Journal Publications

- 1) Frederic Francois, Ning Wang, Klaus Moessner, Stylianos Georgoulas and Ricardo Schmidt, "Leveraging MPLS Backup Paths for Distributed Energy-aware Traffic Engineering," IEEE Transactions on Network and Service Management (TNSM) (under review).
- 2) Frederic Francois, Ning Wang, Klaus Moessner, Stylianos Georgoulas and Ke Xu, "On IGP Link Weight Optimization towards joint Energy-efficiency and Load-balancing," Elsevier Computer Communications (under review).
- 3) Frederic Francois, Ning Wang, Klaus Moessner and Stylianos Georgoulas, "Optimizing Link Sleeping Reconfigurations in ISP Networks with Off-Peak Time Failure Protection," IEEE Transactions on Network and Service Management (TNSM), vol.10, no.2, pp.176-188, June 2013.

## Conference Publications

- 1) F. Francois, N. Wang, K. Moessner, and S. Georgoulas, "Leveraging MPLS Fast ReRoute Paths for Distributed Green Traffic Engineering", in Proc. of 2013 IEEE International Workshop on Quality of Service (IWQoS), June 2013 (Short Paper and Poster Session).
- 2) F. Francois, N. Wang, K. Moessner, S. Georgoulas, and K. Xu, "Green IGP Link Weights for Energy-Efficiency and Load-balancing in IP Backbone Networks", in Proc. of 2013 of IEEE/IFIP Networking, May 2013.
- 3) F. Francois, N. Wang, K. Moessner, and S. Georgoulas, "Optimization for time-driven link sleeping reconfigurations in ISP backbone networks", in Proc. of 2012 IEEE/IFIP Network Operations and Management Symposium (NOMS), April 2012, pp.221-228. **(Best Student Paper award)**.

# Chapter 1

## 1 Introduction

### 1.1 Background and Motivation

Nowadays, many new network services, such as cloud computing and high-definition multimedia applications, are gaining popularity with both private individuals and enterprises. These new network services require the support of extensive computer networks in order to provide a good Quality-of-Experience to their end users. Network operators, including Internet Service Providers (ISPs), are encouraged by both end users and government agencies to deploy more extensive networks in order to meet the needs for better connectivity and traffic capacity that these new network services require to perform well. While these new extensive computer networks bring many benefits to society, they consume a large amount of energy during their operation because they are made up of a large number of links and routers. According to a study in [1], European Telecoms consumed 21.4TWh of power in 2011 and this is expected to increase to 35.8TWh by 2020 if no green networking technologies are implemented. In the USA, it is estimated that the energy consumption of computer networks ranges from 5TWh/year to 24TWh/year [2] depending on which equipment are taken into account. Hence, many stakeholders in the computer network industry have rightly concluded that computer networks consume an excessive amount of energy and are very keen to develop and deploy new methods to reduce the energy footprint of computer networks.

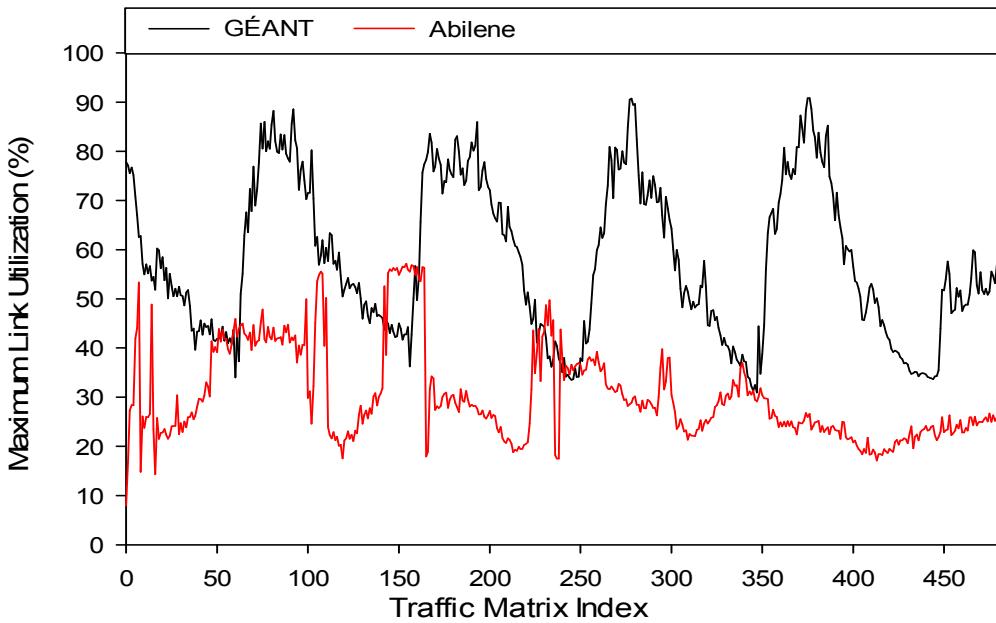
This thesis is primarily concerned with the reduction of energy consumption in wired backbone computer networks through the use of energy-aware traffic engineering schemes. Backbone networks are computer networks which provide the connectivity between the different edge networks to which end users are directly connected. Hence, backbone networks sit at the core of any communication network and when edge networks are upgraded to support more users with larger traffic demand needs, backbone networks need to be upgraded too. It is claimed in [3] that 40% of the total energy consumption of networks will come from backbone networks by 2017 compared to only 10% in 2009. The rationale behind this claim is that large volumes of information will be transferred between different edge networks with the advent of new network services such as cloud computing, Internet-of-Things and high-definition multimedia sharing.

## 1.2 Novel Energy-aware Traffic Engineering Schemes

One important method of improving the energy-efficiency of backbone networks is through the use of Energy-aware Traffic Engineering (ETE) schemes. These ETE schemes route traffic in the backbone networks in such a way that the network devices in the network can either reduce their operating rate or go to sleep completely in order to save energy. ETE schemes can be classified broadly as either offline or online ones.

Offline ETE schemes pre-calculate future energy-efficient network configurations based on historical and/or traffic demand forecasts. The main advantage of offline ETE schemes is the measurement of the traffic demands can be done only once for a long duration of time, i.e. there is no need to constantly monitor the traffic demands in the network, which is a complex operation and adversely affects the overall throughput of the network. Since the calculation of future energy-efficient network configurations is done beforehand, a chosen ETE scheme may be tried with different parameters on a simulated representation of the actual network so as to get the best outcomes. The main drawback of offline ETE schemes is that they can only be applied to well-behaved networks, i.e. if future traffic demands diverge significantly from what is either forecasted or measured before, the pre-calculated network configurations may no longer be optimal. Non-optimal network configurations can cause network instabilities and/or non-optimal energy savings. Figure 1-1 shows the pattern of the actual overall traffic demands in two different backbone networks: GÉANT and Abilene, as measured in [4]. For the five-day period considered in Figure 1-1, the GÉANT network exhibits a predictable/diurnal traffic pattern while the Abilene network has an unpredictable traffic pattern. Moreover, the further analysis of all the traffic demands of the other weeks available in the dataset in [4] shows that GÉANT exhibits a similar regular diurnal traffic pattern over these weeks compared to the first week in Figure 1-1 while the traffic demands of Abilene do not suffer frequently from large variations over these weeks of the dataset. It is not known why Abilene suffers from large variations in traffic demands on a few days of the dataset since no information about the state of the network is available for the period during which the measurements were taken. The possible explanations for the large traffic variations are recurrent failures of the network devices and the bulk transfers of data such as transfers of large scientific datasets between different laboratory locations in USA.

It has already been observed in several operational backbone networks, such as GÉANT [4], AT&T [5] and Sprint [6], that their traffic pattern exhibits a regular diurnal behavior with traffic demands being high during the day and low during the night. Nowadays, most operational networks are over-dimensioned for off-peak time because these networks are dimensioned for peak traffic demands and a limited number of network device failure scenarios such as single link failures.



**Figure 1-1: Predictable and unpredictable traffic patterns.**

Hence, it is possible to put a subset of links in the networks to sleep during the off-peak time in order to save energy while not causing the remaining active links to become overloaded. Even though traffic demands are continuously changing, it is not desirable to frequently optimize the network configuration since this will create instabilities in the network, such as forwarding loops. In this thesis, a novel offline ETE scheme called *Time-driven Link Sleeping* (TLS) is designed to jointly optimize a single energy-efficient network configuration and the off-peak duration for which the configuration will be applied. Hence, only two network configurations are either used for stability reasons: a full network topology and a reduced network topology. A full network topology has all links being active and is used during peak time only while a reduced one has a subset of links put to sleep and is used during off-peak time only.

It is often the case that operational networks suffer from single link failures. Network operators mitigate the adverse effects of such network equipment failure events by overprovisioning their network so that there is enough remaining spare capacity to avoid any packet loss during these events. The TLS scheme reduces the spare capacity in the network during off-peak time by putting links to sleep for energy savings and therefore makes the network more susceptible to packet loss during single link failures. Hence, the second contribution of this thesis is the extension of the original TLS scheme to take into account single link failures. More specifically, a new ETE scheme called *Time-driven Link Sleeping with Single Link Failure Protection* (TLS-SLFP) was designed so that the reduced topology is not only optimized in terms of energy-efficiency performance but also in its ability to prevent any packet loss during single link failures.

Currently, most existing ETE schemes which put links to sleep in backbone networks for energy savings do not re-optimize the link weights. This is not optimal since the original link weights operate in relation with each other and therefore, are only valid when the whole topology is used with no sleeping links. Following this observation, the second major contribution of this thesis is a new ETE scheme called *Green Load-balancing Algorithm* (GLA) which optimizes link weights to enhance the energy-efficiency performance of these existing ETE schemes. In addition to the optimization of link weights for energy savings, GLA can joint optimize the link weights for the more traditional traffic engineering objective of load-balancing. Thus, the main novelty of GLA is its ability to jointly optimize the two essential traffic engineering objectives of load-balancing and energy-efficiency at the same time. At first glance, it is thought that the joint optimization of link weights for both load-balancing and energy-efficiency cannot be done since improved energy-efficiency can only be achieved at the detriment of load-balancing. This is because energy is usually saved by concentrating traffic on the minimum number of links so that the remaining links can go to sleep while load-balancing tries to distribute the traffic as evenly as possible on the maximum number of links in order to reduce the Maximum Link Utilization (MLU) in the network. It will be shown in this thesis that GLA can successfully achieve near-optimal load-balancing while achieving a significant increase in energy-efficiency compared to three existing ETE schemes which do not perform link weight optimization. In addition to improving the load-balancing and energy-efficiency performance, it is important to limit the increase in end-to-end Maximum Packet Delay (MPD) during the operation of ETE schemes. An increase in MPD usually occurs because of the reduction in the connectivity of the networks when links are put to sleep. When the connectivity of a network is reduced, packets have to take longer paths on average and therefore, incur longer delay before reaching their final destination. Unfortunately, most existing ETE schemes do not take into account their adverse effect on the MPD during their operation and therefore in this thesis, the three ETE schemes, which were chosen to evaluate GLA, are modified to take account the MPD. The performance of GLA on these new MPD-aware ETE schemes are evaluated to see the change in performance compared to the original ETE schemes.

Up to now, all the ETE schemes introduced in this thesis were offline ones which can only be applied to operational networks which exhibit predictable traffic demands and therefore, these ETE schemes cannot efficiently deal with unpredictable traffic demands. Hence, the final technical contribution of this thesis is an online distributed ETE scheme called *Green Backup Paths* (GBP). In many operational backbone networks today, Multi-Protocol Label Switching (MPLS) has become a popular intra-domain routing protocol to the detriment of link state routing protocols such as OSPF and IS-IS due to the ease of doing explicit routing [7]. In MPLS-enabled backbone networks, the path between each source-destination pair is signaled as a Label Switched

Path (LSP). In order to provide fast recovery from single link failures, it is necessary to protect each link in all LSPs by installing backup paths. These backup paths consist of links which are already active in the network and therefore, are already consuming energy. Since not all these backup paths are used at the same time, there is an opportunity to re-use these already-installed backup paths to divert traffic away from the protected links and onto the backup links. This enables the protected links to go to sleep since there is no traffic on them. The traffic diverted on the backup paths do not cause the power consumption of the links of the backup paths to increase since these links are already active and do not consume more power when additional traffic is diverted onto them [8]. GBP is a scalable ETE scheme due to its distributed nature where individual routers are concurrently making interference-free routing decisions based on their view of the network. Furthermore, GBP implements a novel interference-avoidance technique which enables routers to make constructive concurrent decisions and therefore, the deterioration of the performance of the network due to conflicting routing decisions is avoided. In addition, this allows GBP to converge faster compared to the scenario where only one decision entity can make decisions at any one time to avoid interference.

The novelties in the three different ETE schemes that are proposed in thesis can be summarized as follows:

- 1) A new ETE scheme called TLS was designed which uses only two optimized network topologies during a day so that instabilities due to frequent network reconfigurations are prevented while still achieving significant energy savings;
- 2) TLS was extended into TLS-SLFP to make the calculated reduced network topology more robust to single link failures while preserving the maximum amount of the energy savings given by the TLS reduced topology;
- 3) A new ETE called GLA was designed to improve the load-balancing and energy-efficiency of existing offline ETE schemes by optimizing the IGP link weights in the network;
- 4) The maximum packet delay performance of existing offline ETE schemes was analyzed and found to be inadequate for the expected Quality of Service that end users expect from the network. Hence, a mechanism was designed to make existing offline ETE schemes aware of the maximum packet delay performance during their operation and keep it within a predetermined bound set by the network operator;
- 5) A new online and distributed ETE scheme called GBP was designed. GBP intelligently re-uses pre-existing backup paths in the network so as to enable links to go to sleep and save energy without affecting the primary purpose of backup paths, i.e. avoid packet loss during single link failures. The re-use of backup paths allows GBP to achieve significant

- energy savings without having to install dedicated paths for its operation and therefore, this results in a lower management overhead of paths for network operators;
- 6) A traffic monitoring mechanism was designed for GBP which allows it to monitor the state of the network and react to sudden traffic changes. Hence, GBP can keep the energy-efficiency of the network optimized even when the traffic demands vary in the network; and
  - 7) An advanced conflict-avoidance mechanism was designed for GBP to enable its deployment in a distributed manner in the network. Hence, GBP is able to converge fast since the different GBP decision entities are able to concurrently make conflict-free decisions.

### **1.3 Thesis Structure**

The thesis is structured as follows:

**Chapter 2:** In this chapter, a complete literature review of the different existing and proposed energy savings techniques for backbone networks is done. The literature review covers the three main categories of energy savings methods in backbone networks: energy-aware traffic engineering, improvement in the energy-efficiency of hardware components and green network protocols which enable network devices to offload their functionality.

**Chapter 3:** This chapter covers the first technical contribution of this thesis which is a new offline ETE scheme called *Time-driven Link Sleeping* (TLS). TLS pre-calculates a reduced network topology and its off-peak period of operation based on a collection of historical traffic matrices. This chapter provides the detailed step-by-step procedure of TLS along with an extensive evaluation of the energy-efficiency performance of the scheme through the use of a Point-of-Presence representation of the European academic network GÉANT and its real traffic matrices.

Furthermore, the TLS scheme introduced earlier is extended to take into account single link failures during off-peak time. This is necessary because the off-peak reduced topology has lower connectivity and spare capacity, and therefore is more susceptible to packet loss during single link failures. A detailed algorithm is given on how the extended TLS scheme, called *Time-driven Link Sleeping with Single Link Failure Protection* (TLS-SLFP), can improve the robustness of the reduced topology obtained from the plain TLS scheme to single link failures so that there is no packet loss and loss of full network connectivity during these failures. This increased robustness should have the minimum amount of impact on the energy-efficiency of the network. Moreover,

the energy-efficiency of TLS-SLFP is evaluated and also compared with that of TLS by using the same network scenario.

**Chapter 4:** This chapter covers the joint optimization of link weights in backbone networks for both load-balancing and energy-efficiency. A meta-heuristic based on a customized multi-objective genetic algorithm called *Green Load-balancing Algorithm* (GLA) is used to solve this challenging problem. GLA is an ETE scheme which can improve already-existing ETE schemes which use link sleeping for energy savings and link state routing for intra-domain routing. Three existing ETE schemes as well as the GÉANT network topology and its real traffic matrices were used to extensively evaluate GLA. In the last part of this chapter, the three chosen ETE schemes are modified so that they take into account the end-to-end MPD performance during their operation. The MPD-aware ETE schemes are then evaluated and compared against their original counterpart.

**Chapter 5:** In contrast with the first two technical chapters, this chapter introduces an online rather than an offline ETE scheme called *Green Backup Paths* (GBP) which is more suitable for operational MPLS-enabled backbone networks which experience frequent unpredictable traffic demands. The chapter provides a detailed description of the mechanism through which GBP can exploit pre-established backup paths to improve the energy-efficiency of the network without adversely impacting the primary purpose of these backup paths, which is providing fast recovery from single link failures with minimal packet loss. Moreover, GBP is a fully-distributed scheme where several routers are able to make concurrent routing decisions that are interference-free. This enables GBP to scale to large networks and also, converge faster. In addition, the effects of GBP on the power and energy saving gains, maximum link utilization, maximum packet delay and change in maximum link utilization and energy-efficiency after the post-failure of single links are evaluated.

**Chapter 6:** This is the final chapter of the thesis where the main contributions of the thesis are summarized but most importantly are put in perspective, i.e. how these different ETE schemes designed in this thesis actually fit in the vast landscape of energy savings methods that are either already used or proposed. In the second part of this chapter, the future lines of research in the field of ETE are briefly discussed.

# Chapter 2

## 2 Literature Review

### 2.1 Introduction

The whole Internet consumed 0.9TWh in 2007, which represented 5.5% of the total energy consumption of the world, and this is expected to increase at the rate of 20-25% per year [9]. European Telecoms consumed 21.4TWh in 2011 and this is expected to increase to 35.8TWh by 2020 if no green networking technologies are implemented [1]. For the USA, it was found out that the energy consumption of computer networks ranges from 5TWh/year to 24TWh/year [2], depending on which equipment are included. These values of power consumption represent a small but sizeable percentage of the overall power consumption in the world. As the user base of the internet increases and more bandwidth-hungry applications become popular, the energy consumption of computer networks will increase rapidly if no green technologies are introduced.

The energy consumption of modern computer networks has become a cause for concern for Internet Service Providers (ISPs). Nowadays, there are five main motivations for reducing the energy consumption of networks. The first one is energy prices are increasing and this is increasing the operational expenditure (OPEX) of network operators. Secondly, most energy sources nowadays are not environment friendly and therefore, networks produce a large amount of greenhouse gases during their operation. The third motivation is that the reliability of devices decreases with increased energy consumption due to increased heat dissipation. The last motivation is the green credentials of network operators can suffer if they are seen as major polluters.

Large ISP networks can be broken down into edge, aggregation and backbone segments. Currently, according to an internal Alcatel-Lucent report quoted in [3], aggregation and backbone segments use around 30% of the total energy consumption of networks while the remaining 70% is consumed by edge segments. Edge segments consume more because they comprise more network devices which include customer premises equipment. Equipment in edge segments are less energy-efficient compared to the professional equipment in aggregation and backbone networks. This is why there has been a drive for more energy-efficient power supplies for customer premises equipment and for this equipment to support different power modes such as

sleep and idle power mode because users of this equipment do not use it 24 hours a day. By 2017, it is claimed that 40% of the total energy consumption of networks will come from backbone segments and therefore, it is worth considering saving energy in this part of the network as well [3].

This chapter is organized as follows: Section 2.2 provides the details about the energy consumption of the different network devices and their components. Section 2.3 puts ETE in the wider context of energy-savings strategies that are/will be available to network operators by providing an overview of the different complementary strategies to ETE, such as energy-efficient hardware and green protocols, which network equipment manufacturers and network operators can leverage in order to save energy in networks. Section 2.4 provides a comparison of the different existing ETE schemes in literature and gives a short description of the inner workings of each scheme. Finally, Section 2.5 gives the conclusions that can be drawn from the analysis that has been done on the various existing ETE schemes.

## **2.2 Energy Consumption of Network Devices and Components**

The way that the total energy consumption of the network is distributed among the different network devices and components can be regarded as a useful guide for researchers to identify where to concentrate their efforts for reducing the energy consumption of networks.

Backbone networks use optical fibers with Wavelength Division Multiplexing (WDM) as the communication medium. It was found that 90% of the energy consumed by the network is by the Internet Protocol (IP) routers with transponders and amplifiers only using about 5% and 3% respectively [10]. This shows that most of the energy consumed by an optical backbone network is located in the IP routers, and researchers should mainly target IP routers in order to save energy. Table 2-1 shows the breakdown of the energy consumption in a Cisco 1200 series router [11], which is a typical router used in backbone networks. Furthermore, it has been demonstrated that the energy used by line cards does not significantly depend on the amount and/or type of traffic passing through the line cards [12]. Therefore, putting line cards to sleep is currently the best option for energy saving in backbone networks but in the future, line cards can have different energy usage characteristics depending on the amount and/or type of traffic being received and forwarded by them [8]. When this type of line card becomes widely deployed, new complementary technologies will have to be designed to make the most out of these new energy characteristics of line cards. For example, it might be advantageous to transmit at maximum speed and save energy by transmitting for a short period of time compared to transmitting at a lower speed for a longer period of time if this mode of operation is compatible with the applications using the line cards. It is likely that line cards which are capable of doing rate adaptation will

become widely-available in the near future since adaptive link rate has already been introduced for Ethernet in the IEEE 802.3az standard for energy-efficiency [13].

### **2.3 Network Energy Saving Strategies**

In recent years, various green technologies have been proposed in the research community towards reducing the energy consumption of future network devices. These green technologies include hardware and protocol modifications. For e.g. the GreenTouch consortium, a group of institutions which aims to improve the power efficiency of networks by 1000 times by 2020 compared to 2010, is investigating the different energy-saving technologies for backbone networks [14].

Improving the energy-efficiency of the hardware components of networks will most likely lead to significant energy savings, but this is considered as a long-term strategy because the introduction of energy-efficient devices will have to be included in the operators' cycle of infrastructure upgrade which involves long time periods. Hardware optimization for improved energy-efficiency can be done through different ways: improvements to circuit design, use of more energy-efficient components, improvements to the layout inside routers for better air flow and support of different power modes. In more detail, the different power modes can be implemented through the use of technologies such as Dynamic Voltage Scaling (DVS) and Dynamic Frequency Scaling (DFS). In [14], it is claimed that improvement in the energy efficiency of hardware components of backbone networks will lead to a 27 times reduction of the energy consumption of the whole backbone network. The support of different power modes in networking equipment can greatly improve the energy-efficiency provided by the different ETE schemes as will be explained in more detail in Section 2.4.

Modification to protocols can be considered as both a short and a long term strategy because it depends on the extent to which protocols are modified to support green objectives, and also how long the standardization of protocols takes. For example, data compression has been proposed in [15] to reduce the amount of traffic in backbone networks and therefore save power by either putting more links to sleep and/or reducing the operating rates of the links.

Another greening strategy is Energy-aware Traffic Engineering (ETE). ETE advocates for the use of existing protocols to change the pattern and amount of traffic in a network so as to maximize the amount of spare capacity that can be removed from the network by putting network elements to sleep. While completely new and heavily-modified network protocols will need a long time for evaluation and implementation by manufacturers of networking equipment, existing protocols can be used for quick implementation of green initiatives in wired networks if only small modifications are made to them.

Even if long-term solutions for greening the future Internet will lead to significant reductions of energy consumption, short term solutions are of equal significance because the environmental issue and the scarcity of energy resources are already regarded as imminent and valid concerns. Hence, ETE is a subject worthy of investigation because of its great potential to significantly reduce the energy consumption in wired networks in the short-term. When new energy-efficient and green protocols are introduced in wired networks at a later time, ETE schemes will still maintain their importance because ETE is complementary to these other energy-savings strategies. Furthermore, ETE schemes can leverage the new capabilities of the energy-efficient hardware and protocols to achieve even better improvement in the energy-efficiency of the network.

## **2.4 Energy-aware Traffic Engineering (ETE)**

ETE schemes attempt to reduce the energy consumption by optimizing the capacity of the network so that there is a better match between traffic demand and network capacity. The rationale is that any excess network capacity leads to too many network devices being fully active and/or working at a higher than necessary rate and therefore, energy is wasted. Moreover, ETE schemes can redistribute traffic in the network in such a way that there is more opportunity to put more network devices to sleep for improved energy-efficiency.

### **2.4.1 ETE Classification**

The different ETE schemes in the literature can be classified through the use of four main criteria. The first criterion is whether the ETE scheme performs changes to the network based on either the historical state of the network or the current state of the network. Offline ETE schemes use the historical state of the network while online ones use the current state to make decisions. The advantage with offline ETE schemes is that the amount of overhead management traffic is low because these schemes use historical traffic demands and/or traffic demand forecasts to make decisions rather than sampling the state of the network continuously.

Furthermore, offline ETE schemes usually do not require large modifications to the way networks operate and therefore, implementation is easier compared to online schemes. On the other hand, online ETE schemes can usually make more accurate and precise decisions because the state of the network is sampled more frequently. Therefore, the potential for energy savings is higher compared to offline schemes. Moreover, Online ETE schemes can react quickly to unexpected events such as link failures and sudden traffic changes and therefore, are more robust compared to offline ones.

**Table 2-1: Breakdown of the energy consumption of a typical backbone router.**

Network Device	Power Used (W)	% of Total Power Used
Line cards	508	42.8
Route Processors	76	6.4
Chassis Components	602	50.8
<b>Total</b>	<b>1186</b>	<b>100</b>

The second criterion by which existing ETE schemes can be categorized is the way they are deployed in a network. The two main ways are: distributed and centralized deployment. In the case of distributed deployment, the decisions are made by many separate entities in the network where each entity usually has only a local view of the network. For centralized deployment, decisions are made by only one central entity in the network. This central entity usually has a global view of the network. The advantages of distributed deployment are that decisions are usually made quicker because information is often gathered from a small number of nearby devices and since the amount of information is small and several decisions entities are operating concurrently, the required processing power and/or time per decision entity is small too.

The advantage of a centralized deployment is that a complete view of the network is used to make decisions and therefore, these decisions are more likely to be globally optimal. In contrast, a decentralized approach may only achieve locally optimal solutions because of the limited view that the separate decision entities have of the network. A decentralized ETE scheme may suffer from degradations in performance due to uncooperative decisions among decision entities. To solve this problem, it is usually required to set up a complex cooperation mechanism between the different decision-making entities. A distributed approach requires several decision entities to be deployed, configured and maintained in the network. In comparison, a centralized approach requires only a single decision entity to be deployed in a network.

The third criterion by which ETE schemes can be categorized is the technology that they use to route traffic inside the network. There are two main varieties of routing technologies which are widely deployed in operational networks nowadays: Internet Protocol (IP) and Multi-Protocol Label Switching (MPLS). Examples of IP protocols are Open Shortest Path First (OSPF) and Intermediate System-Intermediate System (IS-IS). One advantage of IP is that it is a simple routing protocol with no management overhead needed to set up the specific paths compared to MPLS. On the other hand, MPLS can explicitly define which path a Source-Destination (SD) flow will take in the network. This capability of MPLS allows an optimal utilization of all the links of the network to be achieved. The number of paths used to forward each SD flow can be higher when MPLS is used compared to IP because MPLS is not constrained by the link weights as IP is.

Finally, ETE schemes can be categorized according to the hardware technology which they leverage in order to reduce the amount of energy consumption in network. There are two primary hardware technologies: *rate adaptation* and *sleeping*. As mentioned in section 2.2, the energy usage of current technologies employed in backbone networking equipment does not significantly depend on the amount of traffic but this is likely to change in the future. Researchers have already started working on designing new ETE schemes which will exploit link rate adaptation since hardware supporting such a feature is likely to be widely deployed soon [16].

Table 2-3 and 2-4 provide a classification of the major existing ETE schemes according to the criteria just mentioned above.

**Table 2-2: Taxonomy of existing offline ETE schemes.**

ETE Scheme	Characteristics			Performance	
	Deployment	Routing	Technology	Network scenarios for experiments	Main Results
Greedy_Node_Link_Sleeping [17]	Centralized	IP	Sleeping	Virtual network with backbone, aggregation and edge routers in $x$ , $3x$ and $12x$ ratios respectively	50% of routers and 30% of links put asleep
EPAR [8]	Centralized	MPLS	Rate Adaptation	Real network topology with 50 routers and 88 links	60%, 80% and 95% energy savings for logarithmic, linear and cubic energy profile respectively
MILP_WDM [10]	Centralized	MPLS	Rate Adaptation	6 routers and 8 links network, 15 routers and 21 links NSFNET network and 24 routers and 43 links US network	25 to 45% energy savings which increase with network size
MLTE [18]	Distributed	IP	Sleeping & Rate Adaptation	15 routers network	60% energy savings
Green OSPF [19]	Distributed	IP	Sleeping	Altered version of “Exodus” network	60% of links put to sleep
Greedy_Bundled_Link_Sleeping [20]	Centralized	IP	Sleeping	Abilene (12 cables per link), hierarchical (9 cables per link) and Waxman (12 cables per link) networks	83%, 75% and 75% energy savings
GreenTE [11]	Centralized	MPLS	Sleeping	Abilene, GÉANT, Sprint, and AT&T network	27% to 42% energy savings
MILP_Router [21]	Centralized	MPLS	Sleeping & Rate Adaptation	8 routers and 24 links Telstra, and 22 routers and 86 links North American network	12% to 75% and 6% to 61% energy savings for Telstra and North American respectively

*continues on next page*

*continued from previous page*

ETE Scheme	Characteristics			Performance	
	Deployment	Routing	Technology	Network scenarios for experiments	Main Results
Switching_Fabric [22]	Distributed	-	Sleeping	8 links and 8 fabric interfaces per router	6.8 fabric interfaces put to sleep on average
Green_Router [23]	Distributed	MPLS	Rate Adaptation	Abilene network	50% to 88% energy savings depending on maximum delay bound
Connectivity_Sleeping [24]	Centralized	IP	Sleeping	Ebone, Exodus and Abovenet network	74%, 81% and 62% of links put to sleep
Random_Graphs_Sleeping_Rate [12]	Centralized	MPLS	Sleeping & Rate Adaptation	ER, PL, WS network models	Sleep mode efficiency in PL > ER > WS. Cut-off point for rate adaptation was 50% split between fixed and variable energy
DPRA [25]	Distributed	IP	Rate Adaptation	28 routers and 41 links custom network	0% to 25% energy savings depending on traffic load and approximation of cost function
LLN-WS [26][26]	Centralized	IP	Sleeping	8 routers and 28 links custom network	32% to 2% power savings depending on traffic demands
LNF-WS [26]	Centralized	IP	Sleeping	8 routers and 28 links custom network	52% to 4% power savings depending on traffic demands
Green_Dijkstra [27]	Centralized	MPLS	Sleeping	Abilene and NSF networks	4.4% average power savings gap compared to optimal
Green_MST [27]	Centralized	MPLS	Sleeping	Abilene and NSF networks	0.7% average power savings gap compared to optimal
LR-HS & DLFM [50]	Centralized & Distributed	IP	Sleeping	GTE, NSF, EON and USA networks	35% to 50% increase in power savings compared to simulated annealing
MILP-EWO [28]	Centralized	IP	Sleeping	Large number of networks, e.g. Rocketfuel networks and Abilene	12% maximum power savings gap compared to relaxed upper bound
LLLFI [29]	Centralized	MPLS	Sleeping	Large number of network from SNBLib	33% to 50% increase in power savings compared to random link sleeping
QoS-ESIR [19] [49]	Centralized	IP	Sleeping	Ebone, Exodus, Abovenet	Around 50% gain in power savings compared to conventional OSPF routing

**Table 2-3: Taxonomy of existing online ETE schemes.**

ETE Scheme	Characteristics			Performance	
	Deployment	Routing	Technology	Network scenarios in experiments	Main Results
Greening_ Internet [30]	Distributed	MPLS	Sleeping & Rate Adaptation	Simple illustrative network	Highly dependent on inter-packet arrival time
Queue_Rate [31]	Distributed	IP	Rate Adaptation	Peer-to-Peer application running on university network	59% energy savings in links
Compare_Sleeping_Rate [2]	Distributed	-	Sleeping & Rate Adaptation	Abilene network	Sleeping: 78% to 40% time asleep for 2% to 28% utilization. Rate adaptation: 90% to 70% rate reduction for 2% to 28% utilization
EATe [32]	Distributed	MPLS	Sleeping & Rate Adaptation	Abovenet, ATT, Genuity, Sprint, and Tiscali network	2% to 12% energy savings for rate adaptation. 11% to 24% of routers put to sleep. 15% to 30% of links put to sleep
Framework_Rate [33]	Centralized	MPLS	Rate Adaptation	Custom network with QoS and energy model	Mathematical proofs for optimality of proposed algorithms
AWM_EARTH [34]	Distributed	MPLS	Rate Adaptation	Telecom Italia network	Up to 70% energy savings in edge segment of network during off-peak hours
GDRP_PS [35]	Distributed	IP	Sleeping	15 routers and 28 links network	18% energy savings
Bulk_Rate [36]	Distributed	MPLS	Rate Adaptation	500 routers and 2462 links “Molloy and Reed” network	35% energy savings
Transfer_Sleeping [37]	Distributed	MPLS	Sleeping	159 routers and 614 links network, and 244 routers and 1080 links network	Network 1: 44% to 47% energy savings. Network 2: 46% to 49% energy savings
Energy_critical_paths [38]	Centralized	MPLS	Sleeping	GÉANT and Genuity	30% to 40% power savings compared to OSPF for GÉANT. Maximum power savings gap of 10% compared to optimal for Genuity.
GRiDA [39]	Distributed	IP	Sleeping	2 anonymous ISP networks and GÉANT	Around 20% below optimal power savings
Green_darcom [40]	Distributed	IP	Sleeping	GÉANT and Abilene	46% and 22% average gain in power savings compared to plain multi-topology routing
ROD [41]	Centralized	IP	Sleeping	Abilene, Cernet, random	20% power savings over actual routing scenario
Game_LB_ETE [42]	Distributed	MPLS	Rate adaptation	NSFNET	load-balancing inverse to energy savings
LB_ETE [43]	Distributed	MPLS	Rate adaptation	Custom network: 4 ingress and egress routers with 15 backbone routers	Close to optimal power savings for high traffic demands.
LSP_Reroute [44]	Distributed	MPLS	Sleeping	Custom Network	32% to 52% power savings
Green_Cognitive_Network [45]	Distributed	MPLS	Sleeping	Custom Network	6% to 33% power savings

## 2.4.2 ETE Analysis

This subsection provides an in-depth analysis of the different ETE schemes in Table 2-2 and Table 2-3. First, the offline ETE schemes are analyzed followed by the analysis of the online ones.

### 2.4.2.1 Offline ETE Schemes Analysis

#### 1) Sleeping Schemes

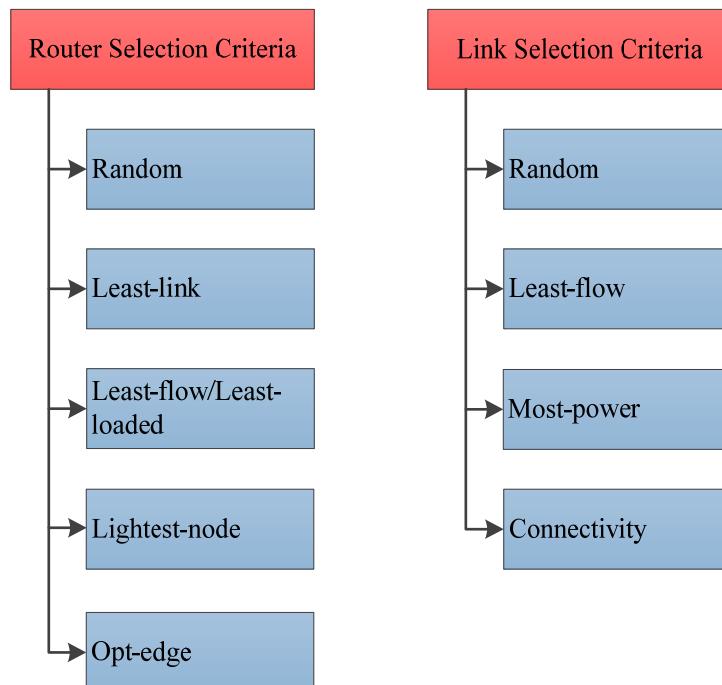
One of the main techniques through which ETE schemes achieve reduction of energy consumption is router and/or link sleeping. This technique is performed within the constraints of maintaining a certain Maximum Link Utilization (MLU) and keeping the full connectivity of the network, i.e. there is always a path between any pair of ingress and egress routers in the network. It has been proven that choosing the optimal set of links and routers to be put to sleep is a NP-hard problem [11] and therefore, there is no computationally-efficient algorithm to solve this problem for any network of practical size. Hence, there is a need to use heuristics to solve this problem efficiently for large sized network scenarios. Research papers in the literature have used different features of the network in order to make an optimized selection of network devices to put to sleep. They believe that their selection method will allow them to obtain a solution that is as close to the optimal solution as possible while using an acceptable amount of time and/or computing power. Figure 2-1 shows the different selection criteria that have been used in the literature for router and link sleeping.

The authors of [17] take a sequential approach where they try to put the maximum number of routers to sleep first and then afterwards, the maximum number of remaining active links. Routers are targeted first because they consume more energy compared to other network devices such as links. If a router is put to sleep, then all the links connected to it is also put to sleep and therefore, a significant amount of energy is saved when a router is put to sleep along with its links. An interesting observation from Table 2-1 is that a router consumes a significant amount of energy even if all the links of that router are put to sleep and therefore, routers must be explicitly put to sleep if possible to save the greatest amount of energy and also make the selection process shorter. Unfortunately, it may not be possible to put a whole router to sleep in some networks, such as GÉANT and Abilene, because all the routers are source routers. In [17], several methods for router selection are compared: random (R), least-link (LK), least-flow (LF) and opt-edge (OE). For the first method of router selection “random”, routers are selected randomly to be put to sleep while for the “least-link” selection method, active routers are selected iteratively to be put to sleep with the active router which has the least number of links connected to it being selected first to be put to sleep. The “least-flow” is similar to the last selection method with the exception that the active router which has the least amount of traffic passing through it being selected first for

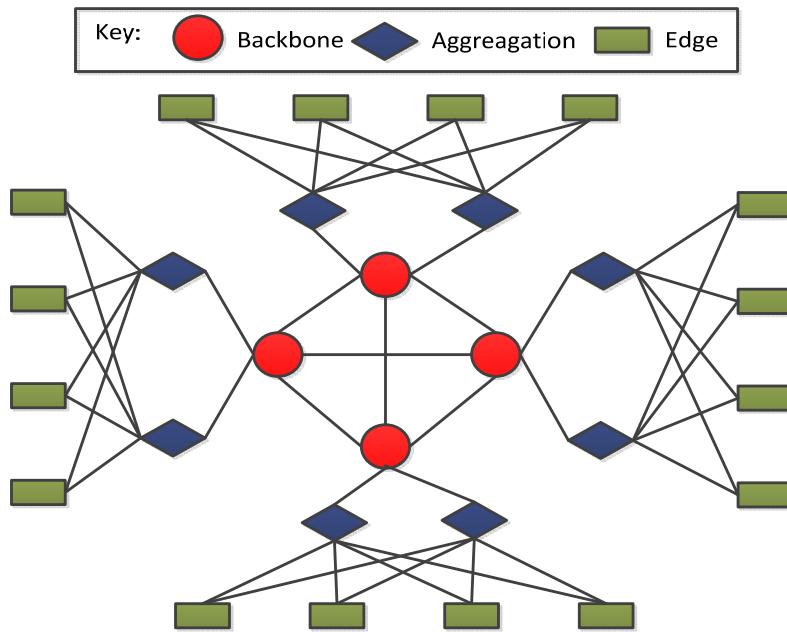
sleeping. Moreover, this paper also uses a special network topology where edge routers are multi-homed to many aggregation routers and therefore, it is possible to put some of the aggregation routers to sleep without the network becoming disconnected. This method is called opt-edge. An example of such network topology is given in Figure 2-2. The methods compared for link selection are: R and LF.

It was found out that the best combination of selection methods for routers and links was OE and LF respectively. In [46], the authors use most power (MP) as an additional method for selecting routers and links to put to sleep. The “most power” method iteratively selects the network device which currently consumes the most amount of power as the first candidate to be put to sleep. They found that MP can be better compared to LF for some specific ISP networks where the edge and aggregation segments of the networks are overprovisioned. In these networks, LF tends to put links in the aggregation and edge segments to sleep first and therefore, LF diverts traffic to the more power-hungry backbone segment of the ISP network.

The authors in [29] develop a new greedy algorithm which also puts the least loaded link to sleep first but with the key difference that for each traffic matrix, the traffic demands are randomly selected and routed along the shortest path. The shortest path is obtained by using link weights which are proportional to the load of the links so that the more a link is loaded, the less likely is it going to be chosen to route a traffic demand. The authors evaluate extensively their algorithm based on several network topologies from the SNBLib collection of network topologies.



**Figure 2-1: Selection criteria for router and link sleeping in [17], [24], [26], [29] and [46].**



**Figure 2-2: Illustrative ISP network with backbone, aggregation and edge segments.**

The research paper [26] compared four router selection schemes: Lightest Node First (LNF), Least Loaded Node (LLN), LNF with Weight Setting (LNF-WS) and LLN with Weight Setting (LLN-WS). LNF ranks routers according to the sum of the capacity of the links connected to them and LLN is similar to LF. It was found that LNF performed better than LLN for many network scenarios with different average traffic demands. The two basic schemes LNF and LLN were improved upon to create two new schemes LNF-WS and LLN-WS respectively. LNF-WS and LLN-WS improve the energy-efficiency of their plain counterpart by optimizing the OSPF link weights after each router removal so that MLU is kept below a set threshold. It was found that LNF-WS and LLN-WS provided a small increase in power savings compared to their basic counterpart for high traffic demands only.

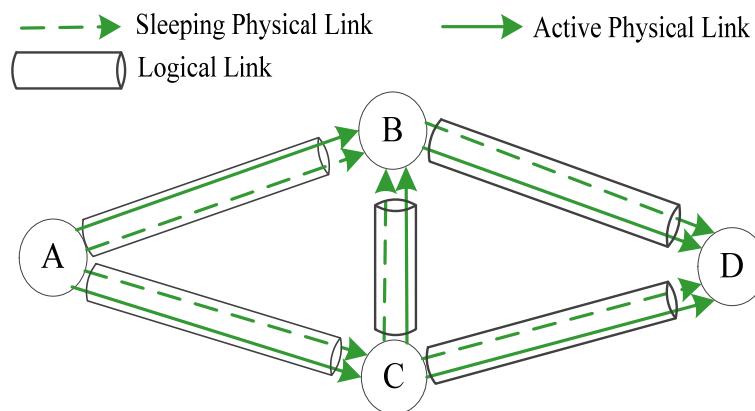
In [24], the algebraic connectivity of the network is used to perform the selection of links to put to sleep. The links which cause the least drop in connectivity are chosen first and the removal of links is stopped once the connectivity of the network drops below the threshold set by the network operator. This method results in networks staying highly connected while reducing their energy consumption.

The authors in [20] use a network model where a link between any two routers in the network is called a logical link and consists of a bundle of physical links. An example of such a network model is shown in Figure 2-3. They then develop three types of greedy link sleeping algorithms for these physical links: Fast Greedy Heuristic (FGH), Exhaustive Greedy Heuristic (EGH) and Bi-level Greedy Heuristic (BGH). FGH removes the least loaded link while EGH removes the least loaded link which will cause the diverted traffic to go onto a new shorter path. BGH is similar to EGH except for the fact that it considers link removal in pairs. The authors found out

that all three algorithms perform similarly in terms of removal of links but FGH took much less time than the other two algorithms, especially BGH.

In addition, the authors in [27] also use the same network model where links are made of a bundle of physical links. They develop two ETE schemes: green Dijkstra and green Minimum Spanning Tree (MST). The green Dijkstra algorithm routes traffic demands iteratively so that the minimum number of physical links is active in the network. Before each routing of one traffic demand, the link weight of each link is updated so it is equal to zero if there is enough spare capacity on the physical link to accommodate the traffic demand. Otherwise, the link weight is set to the value of the power consumption of a physical link if an additional physical link needs to be woken up. Moreover, the authors develop an alternative scheme called green MST where several MSTs are calculated to satisfy all the traffic demands. An additional MST is calculated only if the previous MST cannot satisfy additional traffic demands. Similar to the previous scheme, a MST is calculated based on link weights where a link weight is based on the amount of load on the link and whether the link is either the source or destination router of an already-routed traffic demand. The authors find out that the green Dijkstra and MST schemes perform quite close to the optimal solution with the green MST performing better than green Dijkstra on average. The optimal solution is found by using the “branch and bound” optimization method but this method is only suitable for small networks because the computational complexity of this method is high.

The authors in [11] used a Mixed Integer Linear Programming (MILP) solver instead of a heuristic. A MILP solver can optimally solve the problem of putting the maximum number of links to sleep but it usually takes an unacceptable amount of time to find the optimal solution. In order to alleviate this significant drawback, the authors decide to pre-compute the paths that each SD flow can take. Hence, the MILP solver is limited to choosing which traffic splitting ratio to use for the different alternative paths of the SD flow. This new problem formulation allows for a much shorter computational time but it is not the true optimal solution because only the pre-calculated paths can be used.

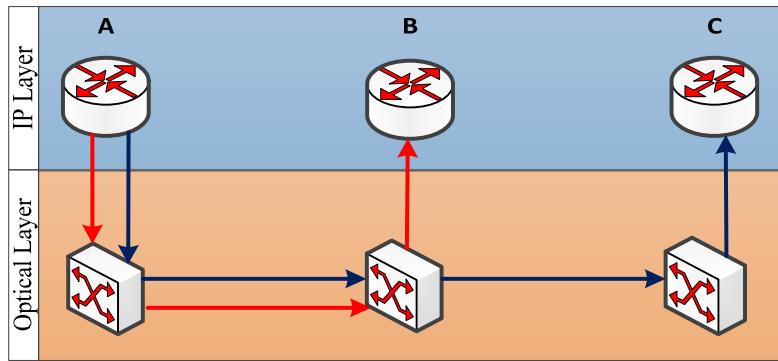


**Figure 2-3: Network where a bundle of physical links connects any two routers.**

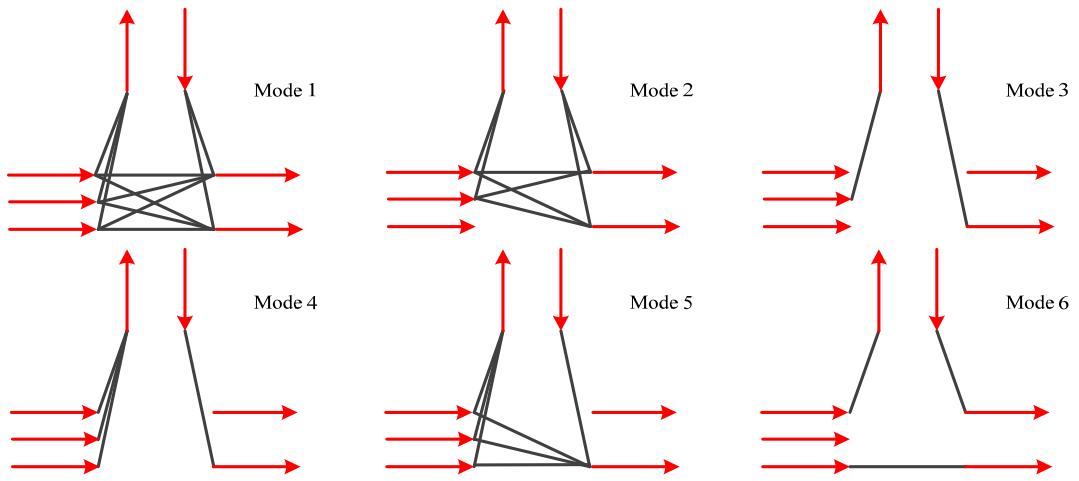
The paths for each SD flow were calculated by three different methods: the  $K$ -shortest paths, the  $K$ -shortest paths which meet a maximum hop count constraint and the  $K$ -shortest paths which meet both a maximum hop count and an end-to-end maximum packet delay constraint.  $K$  is an integer variable chosen by the authors as a trade-off between computational time and path diversity. It was found that while all the different methods for the selection of paths offer significant energy savings, the best method heavily depends on the characteristics of the network topology and in some cases, the amount and pattern of the traffic demands.

In [10], researchers find out that the transformation of data packets from the optical into the electrical domain consumes a large amount of energy. Therefore, researchers investigate the problem of the correct assignment of data packets to a particular wavelength in an optical network so that the chosen wavelength will undergo the minimum number of transformations into the electrical domain for routing and therefore, save energy. This process is illustrated in Figure 2-4 where IP router  $A$  will send packets to destination  $B$  and  $C$  through two different optical channels so that packets destined to  $C$  do not have to be processed by IP router  $B$ . Hence, unnecessary conversion of packets from optical to electrical domain is avoided. This energy-saving method is called optical bypass. Similar to [11], the authors employ a MILP solver in order to find an energy-optimized solution. Moreover, they find out that an energy-optimized solution is also a cost-optimized one since more expensive network devices happen to use more energy. The same observation was made in [47]. Hence, an additional incentive for energy savings in networks is a reduction in the capital expenses due to the use of a smaller number of power-hungry network devices. Moreover, the researchers in [48] consider how to connect the different routers in a given network under the two scenarios of optical bypass and non-bypass so that the total energy consumption of the network is reduced.

In [21], routers are designed to have seven operating modes with each having a different level of energy consumption. Six of the seven operating modes are illustrated in Figure 2-5 with the seventh mode being inactive mode where the router is effectively off and does not forward any traffic. In Figure 2-5, the links at the left and right of each router represent terminating and emanating traffic respectively while the links at the top show local traffic demands. The Mode 1 is active mode where routers function as they would normally. The Mode 2 is links-only mode where the links of the routers can be turned off but the router itself remains switched on. The Mode 3 is bridged-all mode where local demands are bridged to one link and terminating traffic is received on one interface only. The Mode 4 is bridged-local mode where local demands are bridged to one link and terminating traffic is received on all interfaces. The Mode 5 is default-gateway mode where local and transit traffic is forwarded via one link only. The Mode 6 is bridged-many mode where incoming and outgoing links are bridged, including the local link to an outgoing link.



**Figure 2-4: Network where unnecessary optical-to-electrical domain conversions are avoided.**



**Figure 2-5: The different operating modes of a router proposed in [21].**

It was found that for medium to heavily loaded networks, there is not much difference in energy savings for the bridged-all (Mode 3), bridged-local (Mode 4) and default-gateway (Mode 5) modes of operation. In lightly-loaded networks, bridged-all and bridged-local have similar energy consumption but the savings are lower than default gateway. Since default-gateway and bridged-local require greater modifications to the routers, the authors suggest that bridged-all is the mode of operation with the best compromise.

In [49] and [19], routers in the network are classified as either an importer or exporter router. Exporter routers are routers which have high outdegree connectivity while importer routers are the direct neighbors of the exporter routers. The authors modify the conventional OSPF scheme so importer routers use the forwarding table of one of their closest exporter routers. As a result of this modification, importer routers will not use some of their links to route traffic and therefore, these links can be put to sleep. Furthermore, the authors consider QoS by allowing an upper bound to be imposed on the maximum link utilization.

In [50], the authors use Lagrangian Relaxation (LR) to find an initial set of link weights that gives an optimized number of sleeping links for a given network and traffic demands. Afterwards, the

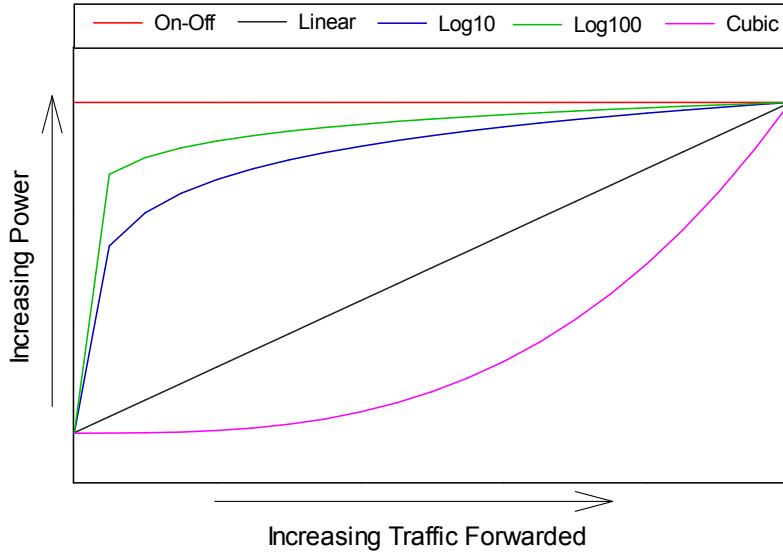
initial set of link weights are fed into an algorithm based on Harmonic Series (HS) to optimize the link weights further. In addition, the authors develop an algorithm called Distributed Loop-Free Routing (DLFR) which is able to update the current link weights in the network to the new ones found by LR-HS without the network suffering from forwarding loops. These forwarding loops may cause the network to suffer from packet loss and link congestion.

Another scheme which does link weight optimization is MILP Energy-aware Weight Optimization (MILP-EWO) in [28] and [51]. MILP-EWO is a complex algorithm that attempts to put to sleep the maximum number of routers and links in the network while satisfying the traffic demands. A combination of the greedy, Link Weight Optimization [52] and near-optimal MILP algorithms is used to make MILP-EWO. MILP-EWO was evaluated on a large set of networks.

## 2) Rate Adaptation Schemes

Even though most network devices currently have an On-Off energy profile, researchers have already explored ways to improve the energy-efficiency of networks where network devices are able to do rate adaptation in order to save energy. In [8], a new ETE scheme called Energy Profile-aware Routing Algorithm (EPRA) is presented where routing is not based on a fixed set of OSPF link weights but the energy consumed by the link. The authors investigated the energy-efficiency of five different energy profiles that current and future links will exhibit as the amount of traffic carried by them varies. The five different energy profiles were: Linear, Log10, Log 100, Cubic and On-Off. These energy profiles are shown in Figure 2-6. According to the authors, the Linear energy profile is exhibited by some switch architectures and the Log10 energy profile is exhibited by equipment which uses hibernation techniques. The Log100 energy profile was used as an energy profile which is mid-way between the On-Off and Log10 energy profile. The Cubic energy profile is the energy profile which is exhibited by equipment using Dynamic Voltage Scaling (DVS) and Dynamic Frequency Scaling (DFS). Finally, the On-Off energy profile is the energy profile exhibited by current equipment. Evaluation of the routing protocol shows that Cubic is the most energy-efficient profile followed by Linear, Log10, Log100 and On-Off. When the traffic demand was scaled up for a network where links exhibit a Cubic energy profile, the energy-aware routing showed an increase in energy savings compared to conventional Open Shortest Path First. It was also observed that energy-aware routing combined with the Cubic energy profile makes the load to be evenly distributed among the links.

Extending the work of [8], [25] proposes a new routing protocol called Dijkstra-based Power-aware Routing Algorithm (DPRA) which is similar to the EPRA in [8] except that rather than doing a global optimization, small fractions of traffic demands are routed using Dijkstra algorithm and the link costs are updated.



**Figure 2-6: Energy profiles of current and future network devices.**

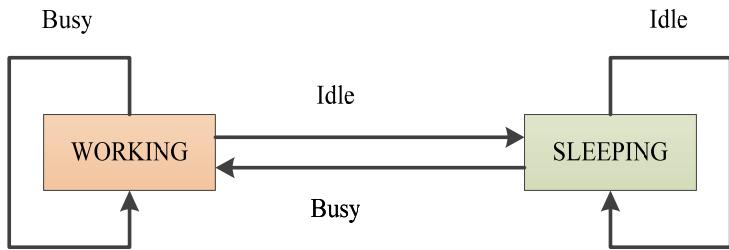
This process is done iteratively until all traffic demands are routed. The advantage of this scheme is that DPRA is quicker to converge than EPRA because EPRA has to resort to the use of approximation of the energy profiles through segmentation in order to have a reasonable computational cost at the expense of lower energy savings.

In [22], the researchers modify the router so that it has an energy consumption which is linearly proportional to the amount of traffic going through the router. In this new router design, the link interfaces will not directly feed into the switching fabric but into an intermediate multiplexer-like component which will merge the link interfaces into a lower number of interfaces to the fabric. This is possible because it is assumed that links are rarely fully utilized and therefore, traffic aggregation can be done. The switching fabric is divided into small sections which can be put to sleep independently. It was found that when the wake-up time is small, the relationship between the number of active interfaces of the switching fabric with traffic volume is mostly linear. As the wake-up time increases, this relationship deteriorates. Hence, the energy-efficiency of routers can be maximized by making the wake-up time as small as possible.

#### 2.4.2.2 Online ETE Schemes Analysis

##### 1) Sleeping Schemes

The use of online sleeping protocols for energy savings was first introduced in [30] where it is necessary for routers to share information so that they can determine which router should go into sleep mode if at all possible during the operation of the network. Unfortunately, the authors did not give the detailed information of how the sleep coordination between routers is performed. A detailed sleep coordination algorithm for routers, called General Routing Protocol for Power Saving (GDRP-PS), is presented in [35].



**Figure 2-7: The two operating states of Power Saving Routers.**

GDRP-PS will run within a backbone segment of the network where routers are not directly connected to end users and therefore, some of these routers can go to sleep without disconnecting an end user. The backbone segment of the network will consist of two types of routers: traditional routers and Power Saving Routers (PSRs).

While traditional routers will run a common routing protocol such as OSPF, PSRs will run GDRP-PS. It is expected that PSRs will consume the same amount of power as traditional routers when active and only 10% of the full power when sleeping. PSRs have two operating states: working and sleeping as shown in Figure 2-7. In a backbone network running GDRP-PS, PSRs will continuously monitor their MLU. Upon detection that their MLU has dropped below a threshold, the PSRs will build a routing table to see if the network will still be fully connected if they go to sleep. If this is the case, they will notify a central PSR. The central PSR is a PSR which has been elected to be the coordinator at the beginning of optimization process. If the central PSR accepts the request of a PSR to go to sleep, the central PSR will send a notification to the PSR so that it can go to sleep. The central PSR replies to requests to go to sleep from other PSRs on a “first-come, first-serve” basis. All PSRs which are not the first PSR to make a request to the central PSR will not get a reply and therefore, cannot go to sleep. The sleeping PSR will wake up at regular interval and use the usual routing protocol to form part of the network again. If the PSR gets a wake-up message within a time interval, it goes into working state and monitors its MLU for the threshold so as to repeat the negotiation process for sleep permission when the MLU threshold is met again. The central PSR will send a wake-up message only if its own MLU is above a certain threshold.

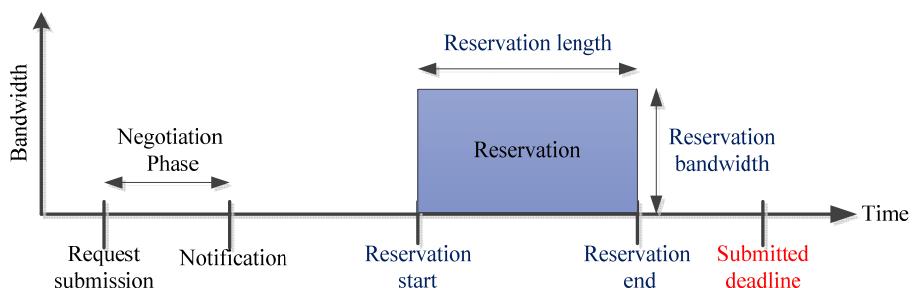
In [36], each router in the network keeps an agenda of the starting time and duration of future bulk transfers of data that will be forwarded by it in the network. Figure 2-8 shows the overall timeline for reserving a timeslot and bandwidth for a bulk transfer of data. An algorithm was then developed for scheduling of the bulk transfer of data where the reservation scheduling process is based on the merged agenda of all the routers. If the merged agenda is clear, then the transfer is scheduled at the middle time between the current time and the deadline if the deadline can be met. Therefore, the algorithm is attempting do the transfer neither too early nor too late. If the transfer is done too early, the transfer will block urgent requests and if done too late, long reservations are

blocked. If the merged agenda is not clear, the algorithm will see where it can fit the transfer so that it happens ideally before the deadline. In this scenario, there may be several options and an end-to-end energy cost model is used to determine what the best option to choose is. An enhancement to the proposed algorithm is that the scheduling can be modified so that transfers happen after each other so that there is no need to switch off-and-on frequently, which is inefficient.

In [38], the authors develop an ETE scheme which is a hybrid between an offline and online scheme. In their scheme, the authors pre-calculate three sets of paths. The first set of paths is to be used when there are low traffic demands in the network. During single link failures, the second set of paths is to be used to avoid packet loss. Finally, the third set of paths is to be used to route the additional traffic demands when the traffic volume increases above a set threshold in the network.

The scheme then uses an online monitoring mechanism to detect changes in the state of the network. After which, the scheme can determine which sets of paths need to be activated. During the evaluation of the scheme, the authors found that the scheme can achieve near-optimal power savings and can also react quickly to single link failures. The optimal solution was found by using a MILP solver but this solver takes an excessive amount of time and therefore it is not a suitable solution for an operational environment but can be used to verify the efficiency of a proposed ETE scheme.

In [39], a distributed online ETE scheme called GRiDA is designed. In GRiDA, each router independently chooses which of its links to put to sleep by using one of many configurations. In each configuration, a different set of links will be sleeping. The decision to choose a particular configuration is usually based on the reward that was obtained when it was chosen before. The reward of a configuration depends on how much the overall power consumption of the network has reduced and if there was any congestion and loss of connectivity in the network when the configuration was used. A router may choose not to use the configuration with the largest reward so as to explore new configurations which may give better reduction in overall power consumption of the network. One drawback of GRiDA is that in the process of finding good configurations, the network may become disconnected and/or congested.



**Figure 2-8: The timeline for reserving a timeslot and bandwidth for a bulk transfer of data.**

In [40], the authors develop an ETE scheme which is based on multi-topology routing and bundle links. The basic idea is to choose one topology for each SD pair in the network so that there are only full and empty physical links in the network as far as possible. The empty physical links can then be put to sleep. The routers in the network have to collaborate to select which one of them will be the first to change the topology that is used to route their SD flows so that there is no conflicting decision.

In [41], the authors design a new algorithm called Routing On Demand (ROD) which optimizes OSPF link weights in the network so that a trade-off between the maximum link utilization and overall power saving is achieved. ROD is based on convex optimization techniques and therefore, the integer problem of putting links to sleep must be relaxed. They found out that the overall power saving decreases when the maximum link utilization decreases.

In [44], the authors reroute label switched paths so that links and their associated line cards can be put to sleep. Each new Label Switched Path (LSP) is found by using the Dijkstra's algorithm where each link in the network uses a V-like function to calculate its link weight depending on the volume of traffic that it is forwarding. The V-like function reduces the probability of empty links being used because this will require an additional link to be activated and more energy to be used. The function also prevents highly-loaded links from being used so as to avoid congestion.

In [31], the authors modify the Transmission Control Protocol (TCP) so that TCP connections can be split in order to allow network devices to go to sleep without terminating the TCP session. The splitting of TCP connections involves adding an extra layer between the socket interface and the application in both the server and client. This layer in the client will instruct its server counterpart to drop a connection when the client goes to sleep and when the client wakes up, its layer will instruct its server counterpart to re-establish the TCP connection by informing it which port to use. If the server wants to send information when the client is asleep, its layer will inform the layer of the client to wake up the client first.

## 2) Rate Adaptation Schemes

In [32], an increasing step function is used to model the power consumption of links with respect to the amount of traffic being forwarded by them. An example of such power profile is given in Figure 2-9. A novel algorithm was developed which performs an optimization of the overall power consumption of the network by removing some load from links which have a power usage near the left edge of a step so that the links use a lower power step. The removed loads will be transferred onto links which have a power usage near the middle of a power step so that an increase in their load will not lead to a higher power step from being used. For example in Figure

2-9, part of the load of link  $L1$  can be moved to the link  $L2$  so that  $L1$  now consumes a lower amount of power  $L1^*$  while  $L2$  still consumes the same amount of power  $L2^*$ .

In [33], the authors create a model of a network where routers and links are both represented as nodes. Each node may have several queues for packets depending on whether it represents a router or a link. The routers have a queue for each QoS packet type while the links have only one queue per link that all the packets on the link must use. The network model makes extensive use of the probability of arrival of data and control packets to determine both the utilization of the routers and also the packet delays. The authors use their network model to build a cost function which is the sum of the power consumed by each node and the packet delays involved. The power model used by a node is similar to the one used in the previous scheme. In order to find an optimal solution to their cost function for the whole network, the authors use a gradient descent method which has polynomial algorithmic complexity and therefore, it takes a lot of time and/or computational power to solve the problem for a large network. Hence, the authors simplified the cost function by integrating the delay metric into the power metric. Another important result of their work is that for their given model of the network, they found out that a network where the traffic is balanced is also an energy-efficient one.

The researchers in [34] use a TCP buffer control called “Active Window Management” (AWM) together with a hardware control algorithm called “Energy Aware service Rate Tuning Handling” (EARTH) in order to achieve energy savings in the edge routers of an Italian ISP network. The authors implement an AWM gateway between TCP sources and destinations. The basic functionality of AWM algorithm is to control the AWM gateway buffers so that they stay at a target level even if the offered load varies. To achieve this functionality, the algorithm calculates a suggested window for each network interface to the gateway and this suggested window is sent to TCP sources connected to that interface so they can adapt their sending rate. The internal function blocks of an AWM gateway is shown in Figure 2-10. EARTH is a hardware control mechanism which tries to increase the capacity of edge routers periodically and check if the AWM can keep its buffer at the required target. If not, EARTH lowers the capacity again. It does this process iteratively to find the optimal capacity to support the network load. It was found that the AWM-EARTH algorithm was able to follow the offered load of the network and keep the gateway buffer at a constant level even when the gateway was not the bottleneck device in the network.

In [2][53][54], “buffering and forwarding” schemes are introduced. These schemes involve upstream routers buffering traffic so that downstream routers do not receive traffic for a period of time and are able to enter sleep mode for an extended period of time [2]. The tradeoff of these schemes is between the amount of buffering and the increased delay that packets experience due to buffering.

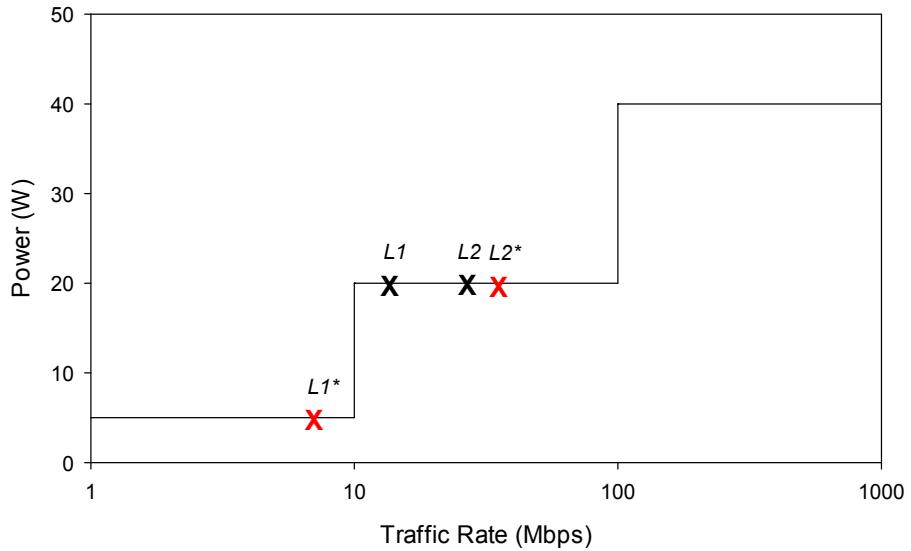


Figure 2-9: Step power profile.

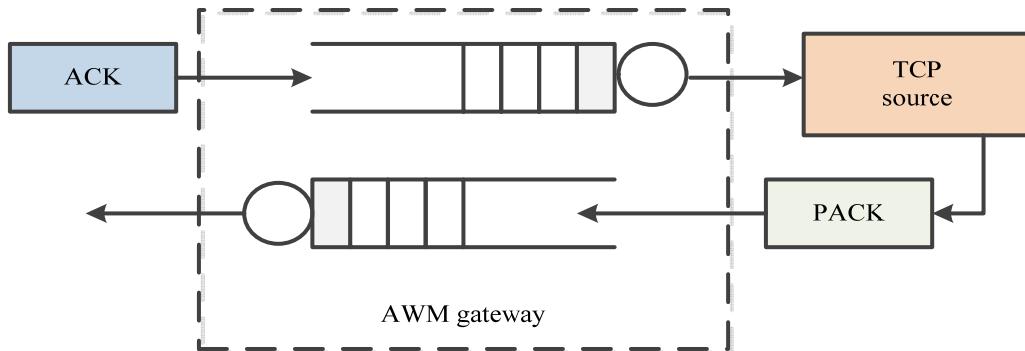


Figure 2-10: Internal function blocks of an AWM gateway.

Another disadvantage of buffering is the upstream routers must have the required amount of memory for storing the buffered packets. Memory with high speed access capability is usually expensive. Another version of this type of ETE schemes is varying the source rate rather than stopping the flow of traffic altogether [53][54]. This energy saving technique implies that the energy consumed by the hardware is significantly dependent on the rate of packet flow in the network. The work in [55] continues along the same line where the effect of different power states of the processing unit of the router on the packet latency and loss rate is investigated. The different power states represent a different tradeoff between the power saved and the time it takes for the processing unit to become active again.

In [2], the authors proposed a distributed rate adaptation algorithm to compare against a “buffer-and-burst” link sleeping algorithm. The rate adaptation algorithm uses either: 4 exponential, 4 uniform or 10 uniform transmission rates. The constraints in this algorithm are that the delay introduced must be below a certain limit and there should not be excessive transitions between the

transmission rates. The algorithm iteratively tries to adjust the transmission rate in order to give the best network performance while attempting not to switch between transmission rates too often. It was found that the “10 uniform transmission rates” solution performed close to the optimum followed by the “4 uniform transmission rates” one. The algorithm was found to have little impact on the delay in the network and can support a transition time of up to 2ms between changes in transmission rates while still having a good performance.

The research paper [37] makes use of a virtual network on top of a physical backbone network. The authors’ solution consists of moving a virtual link away from its original physical link onto a new physical link hosted by a different line card so that the original line card can be put to sleep. The line card selection process is based on a greedy algorithm which uses the “least load” metric as decision criterion. The transfer process will happen at Layer 2 of the Open Systems Interconnection (OSI) model and will be transparent from Layer 3 of the OSI model. A network control unit will trigger the transfer process when the traffic load falls below some predetermined thresholds. Practical implementation of their routing algorithm shows that there is no packet loss and that the throughput of the routers remained the same during the transfer process.

In [42], the authors employ a game theoretic approach to the problem of jointly optimizing the load-balancing and energy-efficiency in the network. They use a polynomial function to model the energy consumption of a link with traffic flowing on it. Therefore, the problem becomes convex and easier to solve. They found the same inverse relationship between load-balancing and energy-efficiency as in [41].

In [45], the authors use a cognitive network where each router broadcasts smart packets to gather the state of the network. When these smart packets return back to their sending router, the sending router can then make a decision based on the information in the smart packets on how to route the data packets in an energy-efficient manner. The authors also use reinforcement learning to enhance their decision-making on how to route packets.

#### 2.4.2.3 Sleeping versus Rate Adaptation

This section investigates how network design choices can affect whether sleeping or rate adaptation is better at energy savings. The first network design choice studied is the connectivity of a network. The three network connectivity types studied in [12] are: Erdos and Renyi (ER), Power Law (PL) and Watts-Strogatz (WS). ER creates links between any two routers based on a given probability and the degree of the routers will follow a Poisson distribution. The ER model exhibits a small-world property and the diameter of the graph is the log of the number of routers. The PL model exhibits a power-law distribution for its connectivity degree between the routers where some routers will act as hubs and have more connections than others. The WS model is first constructed with each router connected to a fixed number of neighbors and then additional

links are added to randomly chosen pairs of routers. This model exhibits small-world and local clustering properties. The performance of sleep mode is higher in PL compared to ER. WS displayed the worst performance. The breakeven point between the energy-efficiency performance of sleep mode and rate adaptation occurs when the ratio of fixed to proportional part of power consumption is equal to one for all models. Sleep mode effectiveness decreases when the number of routers increases and increases when the router degree increases for all three models. The ER and PL models have better path length properties compared to WS. While the study done in [12] uses randomly generated network topologies, the study still has practical significance since network operators can get an approximation of the potential energy savings of their network by comparing their network to the three types of network model described in the study. Moreover, it is possible for network operators to make their network become more energy efficient by modifying their network so that it has similar connectivity characteristics as the PL network model, the most energy-efficient network model in the study.

Future ETE schemes will have to consider the different energy profiles of the devices in a network, especially the energy consumption relationship between the devices. This will, most likely, make it complicated to converge to a solution quickly. Manufacturers of network devices will also have to design a protocol which will help disseminate the information about the energy profiles of their devices so that a “plug n play” approach can be taken with addition and optimization of new network devices to a network.

## 2.5 Conclusions

From the analysis of the different existing ETE schemes in literature, several gaps have been identified which the ETE schemes designed in this thesis have tried to rectify. The first gap is all existing offline ETE schemes presented up-to-now calculate a new network configuration for each traffic matrix and therefore, several network configurations are needed for the energy-efficient operation of the network since the traffic demands vary during the day. This method of operation is not efficient because it introduces the need to switch between different network configurations frequently and this causes the network to become unstable [37]. Hence, an obvious extension to existing offline ETE schemes is the calculation of a network configuration which will jointly offer a significant amount of energy savings and can be applied for a long time in the network so that there is no instability due to frequent reconfigurations. The *Time-driven Link Sleeping* (TLS) scheme, which will be presented in Chapter 3, is an ETE scheme which attempts to solve this problem by using only two network topologies during a day so as to avoid frequent reconfigurations.

Furthermore, it is well-known that backbone networks can often suffer from single link failures [56] and therefore, it is important that any future ETE schemes are robust to these failure scenarios since these schemes make the network more vulnerable to traffic congestion during single link failures because they reduce the amount of spare capacity in the network by putting links to sleep for energy savings. Unlike the majority of existing ETE schemes, the TLS and GBP ETE schemes designed in this thesis take into account single link failures in their operation so as to minimize the probability of packet loss during these failure scenarios.

Another gap in literature is most existing offline ETE schemes do not take into account the traditional traffic engineering objective of load-balancing during their operation. Network operators currently place more priority on load-balancing compared to energy-efficiency because load-balancing makes their network more resilient to traffic congestion during traffic upsurges and unlike energy consumption, traffic congestion is a network performance parameter which can directly affect the customers of the network operators. Therefore, the *Green Load-balancing Algorithm* (GLA), which will be presented in Chapter 4, enables these existing ETE schemes to achieve near-optimal load-balancing while significantly improving the energy savings that they provide. Hence GLA makes it more likely that these ETE schemes will be adopted by network operators since they can improve the energy-efficiency of their network without sacrificing the “more” important objective of load-balancing. The TLS, TLS-SLFP and GLA ETE schemes are offline ETE schemes, which are the most suitable type of ETE schemes for backbone networks which experience predictable/diurnal traffic patterns because of the low complexity and ease of implementation of such schemes.

Not all networks exhibit a regular traffic pattern and therefore, it is important to also design an online ETE scheme to cater for these networks. Most existing online ETE schemes are too complex and do not offer a large amount of flexibility in the way in which they are operated. The *Green Backup Paths* (GBP) designed in this thesis extensively re-uses the existing capabilities of already-deployed networks such as existing backup paths to enable links to go to sleep and Traffic Engineering Link State Advertisements (TE-LSAs) to distribute information about the state of the network to all ETE decision entities. Therefore, GBP does not incur the significant cost and complexity associated with existing online ETE schemes in literature and these features of GBP will encourage network operators to deploy GBP instead of other competing ETE schemes. Moreover, GBP is aware that the logical link between any two routers in a backbone network can consist of a bundle of physical links and therefore, it is able to put some physical links in a bundled logical link to sleep to achieve greater energy savings since it is not restricted to putting a whole logical link to sleep. In order to provide a fast and accurate reaction to sudden traffic changes, an online ETE scheme must be deployed in a distributed rather than a centralized manner and must be able to concurrently make conflict-free decisions. Unlike most existing

online ETE scheme, GBP has an advanced conflict-avoidance mechanism which enables all the GBP decision entities to concurrently make conflict-free decisions. The conflict-avoidance mechanism is also a simple mechanism which does not require a large amount of resources to properly operate.

Moreover, all the ETE schemes designed in this thesis respects five important constraints that all effective ETE schemes should do but many ETE schemes in literature don't. The first constraint is the flow conservation in the network must be conserved, i.e. all the traffic demands should be satisfied even if some network elements have been put to sleep to save energy. The second constraint is any ETE scheme is not allowed to make the maximum link utilization in the network to rise above a set threshold as predetermined by the network operator. This constraint allows the network to have enough remaining spare capacity so as to tolerate sudden upsurges in traffic demands without any packet loss. The third constraint is the ETE should keep the network fully-connected during its operation by ensuring that there is a communication path between any two routers in the network during the normal operation of the network. The fourth constraint is the network should be resilient to single link failures by remaining fully-connected during single link failures, i.e. there must be a communication path between any two routers in the network even if a link fails in the network. Furthermore, the remaining network capacity after a failure has happened must be sufficient to allow all traffic demands to be satisfied without any packet loss. The fifth constraint is Quality of Service should be taken into account during the operation of the ETE scheme by using a single path between each Source-Destination (SD) pair in the network to forward traffic. In addition, the maximum packet delay should be kept below a predetermined threshold during the operation of the new ETE scheme despite the fact that link sleeping reduces the connectivity of the network and this can result in excessively long SD paths to satisfy the traffic demands, if not managed properly.

## Chapter 3

# 3 Optimization of Time-driven Link Sleeping Reconfigurations in ISP Networks<sup>1</sup>

## 3.1 Introduction

Most offline Energy-aware Traffic Engineering (ETE) schemes ([8][10][17][18][19][20][24]) calculate a new energy-efficient network topology, i.e. a new network configuration every 15 minutes typically, for each new traffic matrix (TM). Frequent re-configurations of the network lead to instabilities in the network because it takes time for the forwarding table of routers to converge to a coherent view of the network after the network has been reconfigured [57]. Hence, these ETE schemes present an unacceptable trade-off between energy-efficiency and network stability for network operators to be willing to adopt these ETE schemes. In this view, a new ETE scheme called *Time-driven Link Sleeping* (TLS) is proposed in this chapter. TLS achieves improved energy-efficiency of the network through the use of two network topologies, a full and a reduced one, which are operated in a time-driven fashion. The full network topology is applied during peak time where traffic demands are high, so all the links of the network need to be active to provide the maximum capacity to route traffic and avoid congestion. On the other hand, the reduced network topology is applied during off-peak time where traffic demands are low. The reduced topology contains some links which are put to sleep to improve the energy-efficiency of the network while respecting some constraints such as Maximum Link Utilization (MLU). The main novelty of TLS is its ability to jointly optimize a reduced network topology and its period of operation during off-peak time for energy savings. Since TLS uses only two network topologies, i.e. network configurations, it drastically reduces the instabilities in the network compared to the other existing offline ETE schemes.

---

<sup>1</sup> The content of this chapter is based on the following two already-published works:

1. F. Francois, N. Wang, K. Moessner, and S. Georgoulas, "Optimization for time-driven link sleeping reconfigurations in ISP backbone networks", in Proc. of 2012 IEEE/IFIP Network Operations and Management Symposium (NOMS), April 2012, pp.221-228.
2. Frederic Francois, Ning Wang, Klaus Moessner and Stylianos Georgoulas, "Optimizing Link Sleeping Reconfigurations in ISP Networks with Off-Peak Time Failure Protection," IEEE Transactions on Network and Service Management (TNSM), vol.10, no.2, pp.176-188, June 2013.

The first part of this chapter is organized in five sections as follows: Section 3.2 provides a brief illustration of the mechanism of TLS. In Section 3.3, the problem formulation for finding an optimized reduced network topology and its off-peak period of operation is presented. The detailed operation of TLS is then presented in Section 3.4. The results of the extensive evaluation of TLS on an operational European academic network, GÉANT, and its real traffic matrices are shown and discussed in detail in Section 3.5. In the second part of this chapter (Section 3.6 to 3.10), TLS is extended to take into account single link failures by developing a new ETE scheme called Time-drive Link Sleeping with Single Link Failure Protection (TLS-SLFP). The motivation for developing TLS-SLFP is described in Section 3.6.

## 3.2 Time-driven Link Sleeping (TLS) Overview

As mentioned in the previous section, TLS applies two network topologies in a time-driven approach. The rationale behind this is that many ISP networks have *regular/predictive* traffic behavior patterns ([5][37][58]) which offers the opportunity for simple network topology control rather than relying on complicated online ETE schemes. The proposed algorithm aims to compute a reduced network topology which will give the optimized *combination of the number of sleeping links and duration of the sleeping period*. This energy saving approach can be used because networks are currently dimensioned and operated by taking into consideration traffic demands during peak hours and are further overprovisioned to account for unexpected events. Therefore, there is room for reducing the network topology when traffic demands are low during the off-peak time and save energy while still maintaining the same degree of over-provisioning that occurs during peak time network operation. In order to obtain a robust off-peak network reconfiguration, the calculation of the off-peak topology and its duration is based on multiple sampled traffic matrices as input, and the detailed description of the algorithm will be given in sections 3.3 and 3.4.

According to the proposed scheme, two distinct network topologies are used in daily operation. During the normal operation hours, the original *full topology* is applied for handling customer traffic, as it happens in common practice. During off-peak time, the *reduced network topology* that excludes scheduled sleeping links is used for energy saving purposes. In order to seamlessly perform topology switching without incurring routing disruption, the use of existing multi-topology routing protocols, such as MT-OSPF [59] and MT-ISIS [60], is proposed as the underlying routing platform. Specifically, two routing topologies are configured, one with the full physical topology and the other excluding the links scheduled to sleep. Once the operation has entered the scheduled off-peak time, individual routers simultaneously activate the reduced routing topology by remarking the multi-topology identifier (MT-ID) of packets from the default

full topology MT-ID to the reduced topology MT-ID. Similarly, when the scheduled off-peak time expires, all routers remark customer packets back to the full topology MT-ID.

Such operation avoids the rigid routing re-convergence procedure based on one single topology, which is generally considered to be harmful. From an optimality point of view, a technical challenge of having these two network topologies is that the reduced network topology should not be exclusively optimized with respect to the number of sleeping links, but it should also support as long operational periods as possible with this reduced network capacity, but without causing network congestion incurred due to customer traffic dynamics. It is not difficult to infer that there is a tradeoff which needs to be optimized, between the *number* of sleeping links in the reduced topology and the *duration* of the off-peak time when this topology is actually applied.

This is because, if an excessive number of links is excluded from the network based on a purely greedy approach, then the resulting reduced topology is not able to handle even minor traffic increases, and in this case the full topology has to be restored, which leads to a *very short duration* for the use of the reduced topology.

In order to determine the optimized *combination* of the reduced topology and its applied operation duration, it is important to consider how traffic behavior patterns can play a role. As such, network monitoring is periodically performed which produces distinct traffic matrices at regular time intervals (e.g. every tens of minutes), as is the case in operational networks (GÉANT [4] and Abilene networks [61]). It should be noted that network configurations based on one single traffic matrix are not sufficiently robust in dealing with traffic dynamics, as has been indicated in conventional traffic engineering schemes [57]. Therefore, multiple traffic matrices are considered as input in order to produce the optimized combination of the reduced topology and its off-peak operation duration. To ease the *time-driven* reconfiguration operations, it is also desired that the off-peak configuration starts at exactly the same time and has the same duration on a daily basis (e.g. 7:00PM–7:00AM). According to the proposed solution, a single synthetic traffic matrix is computed based on multiple sampled traffic matrices across a given period (e.g. weekly). These selected traffic matrices represent the traffic pattern during the daily off-peak period. Thereafter, a reduced network topology with the excluded sleeping links is computed based on the single synthetic traffic matrix. Similarly, determining the off-peak duration on a daily basis also takes into account traffic matrix patterns during each day. It is worth mentioning that, although traffic dynamics behaviors exhibit some similar patterns on a daily basis, it is also important to make sure the same off-peak topology can be applied in a unified time-driven manner even though there can be traffic pattern variations at the same (off-peak) time in each day. Input traffic matrices to the algorithm can also be scaled up so as to enhance the robustness of the calculated reduced network topology to potential changes in the traffic patterns. In the next section, the overall optimization problem will be first formulated, followed by the description of the joint schemes of

computing the reduced topology and determining its duration. The outcome is an optimized network topology coupled with a unified time window for its configuration on a daily basis.

### 3.3 Problem Formulation

**Table 3-1: Definition of symbols for TLS.**

Variable	Description
$A_x$	Set of sleeping links in the reduced topology during off-peak time
$\alpha$	Maximum allowable utilization of link capacity
B	Traffic scaling factor
$c_{ij}$	Bandwidth capacity of link from router $i$ to $j$
$f_{ij}^{sd}$	Traffic demand from $s$ to $d$ that traverses link from router $i$ to $j$
$f_{ij}$	Total traffic demand on link from router $i$ to $j$
$G(R, L)$	Directed graph with $R$ being set of routers and $L$ being set of links
$I_x$	Set of consecutive traffic matrices supported by reduced topology $A_x$
$\Pi(t^{sd})$	Set of links taken by $t^{sd}$ from router $s$ to $d$
$t^{sd}$	Traffic demand from router $s$ to $d$

The overall problem formulation on optimizing the reduced topology and its duration can be expressed as:

$$\text{maximize } \max \{|I_x| \times |A_x|\} \quad (3.1)$$

subject to:

$$\sum_{j=1}^{|R|} f_{ij}^{sd} - \sum_{j=1}^{|R|} f_{ji}^{sd} = \begin{cases} \beta t^{sd} & \forall s, d, i = s \\ -\beta t^{sd} & \forall s, d, i = d \\ 0 & \forall s, d, i \neq s, d \end{cases} \quad (3.2)$$

$$f_{ij} < \frac{\alpha}{100} c_{ij} \quad \forall i, j \text{ with } \alpha \in [0, 100] \quad (3.3)$$

$$\Pi(t^{sd}) \cap A_x = \emptyset \quad \forall s, d, x \quad (3.4)$$

Equation (3.1) is the objective function which jointly optimizes the reduced topology (in terms of the number of sleeping links) and the duration of its configuration (in terms of the number of consecutive traffic matrices which can be supported by that topology). The overall off-peak time duration can be calculated from the number of traffic matrices, which are captured at regular intervals, covered by the period. Equation (3.2) is the standard flow conservation constraint. Equation (3.3) ensures that all links have utilization below a given threshold determined by the ISP. That is, with the reduced topology, the maximum link load should not exceed the threshold  $\alpha$  (in terms of the percentage of the link capacity) at any time during the off-peak time. Equation (3.4) makes sure that only the active links can carry the traffic during the off-peak time period. In addition, it is also required that the reduced network topology should not be broken due to the link

removals, and also on a daily basis the starting/ending time for the off-peak topology configuration should be unified.

It should be noted that the simplified problem of identifying the optimal number of links for sleeping based on a single traffic matrix has been proven to be NP-hard [32][30]. Therefore, there is no known computationally-efficient approach to optimally solve the problem and heuristics need to be used. Next, the proposed TLS scheme, which is based on heuristics, is proposed below.

### 3.4 Proposed Heuristic Scheme

Figure 3-1 illustrates the overall approach. The curve in the figure indicates the actual MLU dynamic pattern across a given period (5 days<sup>2</sup>). Such a MLU curve is effectively plotted according to the monitored traffic matrices at a certain time interval. In this example, this is illustrated with the TM measurements taking place every 2 hours, 12 TMs are produced within each 24 hour period. In Figure 3-1,  $M_{v,q}$  indicates the  $q^{\text{th}}$  TM in day  $v$ . Such historical TM information is used for computing future off-peak network configurations and this is analogous to the traffic matrix forecasts used in general traffic engineering. First of all, multiple “sampled” TMs from different days in the considered period are used for computing a synthetic off-peak TM, based on which sleeping links are initially identified in the reduced network topology. The purpose of using multiple TMs instead of a single one is for robustness reasons, as traffic conditions even at the same time point in different days may vary. In addition, in order to compute the actual off-peak configuration period (time window) on a daily basis, a *unified* starting point for computation on each day is be identified, i.e. this point is located at the same  $q$  for every day  $v$ . Starting from this point, an algorithm is designed for expanding (in both directions) the off-peak configuration window size incrementally in order to maximize the overall energy savings based on the reduced off-peak topology. As the final result, the reduced topology can be configured in a unified way on a daily basis, for instance between traffic matrix  $M_{v,q1}$  and  $M_{m,q2}$  within each  $v^{\text{th}}$  day. For instance, in Figure 3-1 the values of  $q1$  and  $q2$  for identifying the unified daily off-peak time are 4 and 7 respectively.

The overall heuristic-based approach consists of the following three sequential stages: (1) computing the synthetic traffic matrix and the starting point for off-peak window size expansion, (2) greedy link removal for the provisionally reduced topology, and (3) the joint expansion of the off-peak window size and the finalization of the reduced topology to be applied during the off-peak period.

---

<sup>2</sup>To better illustrate the daily off-peak configuration windows, the scheme was made to start from midday on the first day with peak traffic volume in the network. Moreover, it is not necessary for the peak/valley MLU to occur exactly at the same time point on each day. This figure is only for illustration purposes.

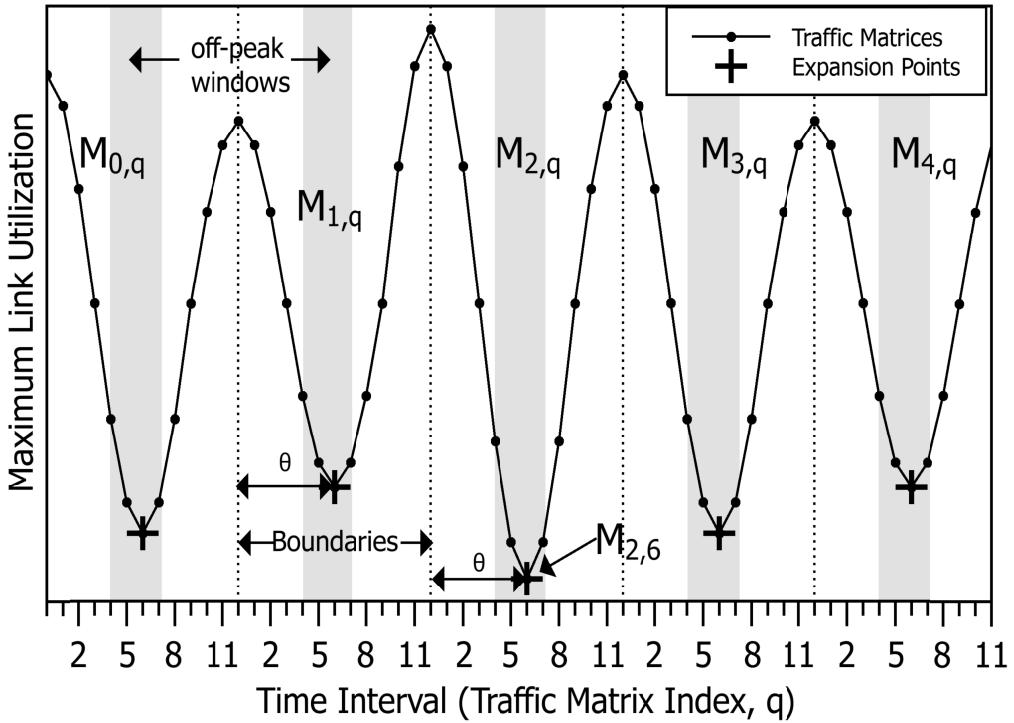


Figure 3-1: Illustration of the overall approach.

### 3.4.1 Stage 1: Computation of the synthetic TM and the starting point for off-peak window size expansion

As previously indicated, the off-peak operational window size can be represented with a sequence of TMs at regular intervals, for instance every 2 hours as shown in Figure 3-1. The MLU of each traffic matrix is calculated to identify the traffic matrix with the *lowest* MLU across the whole period (denoted as  $M^*$ ) together with its “location” in the sequence of the TMs (e.g. TM  $M_{2,6}$  in Figure 3-1).

The 2<sup>nd</sup> index of this traffic matrix with the lowest MLU (i.e. 6 in this specific example) is set to be the *expansion point* of the daily off-peak window. To enable a unified daily off-peak window, the location of the expansion point in every other day is at the same position as  $M^*$ . This is represented in Figure 3-1 through the calculation of distance  $\theta$  from the beginning of each day. Put in other words, since  $M^*$  in the figure is  $M_{2,6}$ , then the expansion point in other days should be  $M_{v,6}$  (i.e.  $v = 6$ ), where  $v$  represents the index of the days under consideration. It is worth mentioning that it is not assumed that the actual lowest MLU always takes place exactly at the expansion point in the other days. But this is not exhibited in Figure 3-1 simply for simplicity reasons. These expansion points are the starting points that will be used in the 3<sup>rd</sup> stage to determine the actual location and the size of the off-peak windows.

With regard to the traffic matrix information to be used for computing the reduced topology with link removal, it is not efficient if only one single sampled TM is used as input, say  $M^*$  with the lowest MLU across the whole period. A more robust strategy is to take into account multiple TMs in order to obtain an optimized network configuration across all scenarios during the period. As such, a synthetic traffic matrix is first computed based on individual TMs located on the daily expansion points. In this example, the overall synthetic TM is computed according to the 5 TMs shown in Figure 3-1. For each traffic demand in the synthetic TM, the value is the average of the corresponding entries from the 5 input TMs at the daily expansion points.

The algorithmic complexity of calculating all the shortest path of all the Source-Destination pairs in a graph and assigning all the traffic demands of one traffic matrix to the links of the graph is given by  $O(|R|^2 \cdot (|L| + |R| \cdot \log|R|))$  based on Dijkstra's shortest path algorithm. Since this procedure has to be done for each traffic matrix and there are  $Q$  traffic matrices in one day and in total  $V$  days are considered as input, the overall algorithmic complexity to determine the whole set of MLUs is  $O(Q \cdot V \cdot |R|^2 \cdot (|L| + |R| \cdot \log|R|))$ . This is actually the overall algorithmic complexity for this stage since the algorithmic complexity of finding the expansion points and calculating the synthetic traffic matrices does not dominate the overall algorithmic complexity.

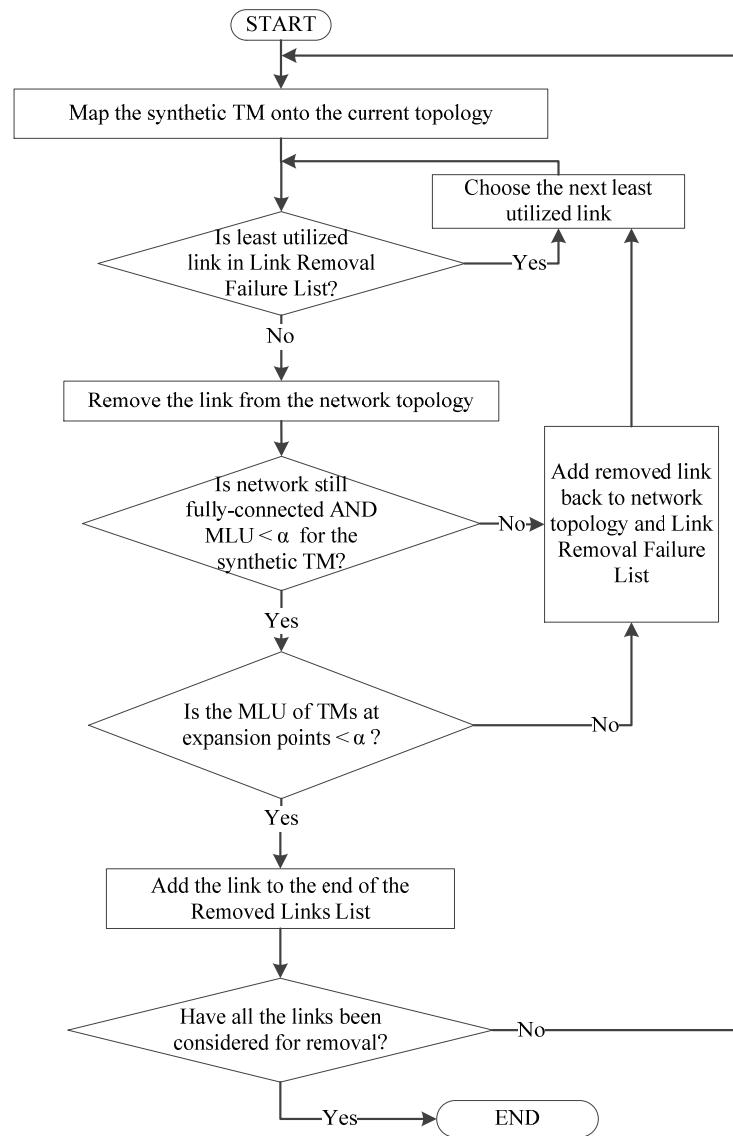
### 3.4.2 Stage 2: Greedy link removal

In this stage, a reduced network topology is provisionally computed based on the synthetic TM obtained from the previous stage. Further fine-tuning of this topology will be performed jointly with the determination of the actual off-peak operational window size in the next stage. The computation of this provisional reduced topology is achieved by the iterative greedy removal of the links from the full network topology. The detailed operation of the link removal in this stage is shown in Figure 3-2.

Specifically, upon the actual removal of any link, the corresponding utilization of all the remaining links is updated in the whole residual topology. If the link under consideration cannot be removed, it is put into the *Link Removal Failure List* so that it will not be further considered. The condition for a successful link removal is that the residual network topology still remains connected and also the resulting MLU does not exceed the predefined constraint  $\alpha$  based on the input synthetic TM. In addition, considering the requirement that the daily off-peak windows need to have unified starting time and duration, for robustness concerns the traffic matrices at all expansion points are also tested for compliance with the constraints in Eq. (3.3) and (3.4) (shown in section 3.3) each time one link is considered for removal from the network topology. At the end of this operation, the network topology *with the least remaining active links*, which still satisfies

Eq. (3.3) and (3.4), is determined as the reduced topology. This topology will be further refined jointly with the optimization of the off-peak operational window size in the next stage.

For each candidate link for removal, *Stage 2* has to find the alternative shortest paths of all the Source-Destination pairs of the graph and route all the traffic demands of the synthetic traffic matrix in order to check for violation of the MLU constraint. This step of *Stage 2* has algorithmic complexity of  $O(|L| \cdot |R|^2 \cdot (|L| + |R| \cdot \log|R|))$  given that maximum  $|L|$  links need to be considered for removal. The algorithmic complexity of checking for full connectivity of the graph after each link removal is  $O(|R| \cdot (|L| + |R| \cdot \log|R|))$  based on router visiting using shortest path algorithm. Since the algorithmic complexity for connectivity check is lower than the previous algorithmic complexity, the overall algorithmic complexity of *Stage 2* is  $O(|L| \cdot |R|^2 \cdot (|L| + |R| \cdot \log|R|))$ .



**Figure 3-2: Flow chart for greedy link removal.**

### 3.4.3 Stage 3: Determination of off-peak window size and the final reduced topology

The final and main stage of the algorithm is to compute the actual off-peak window size jointly with the fine-tuning of the number of sleeping links. The flow chart in Figure 3-3 shows the detailed operation in this stage. To start, the maximum number of *consecutive* TMs which can be supported by the reduced network topology computed in the previous stage is identified. The expansion of the daily off-peak windows starts from the expansion points obtained in *Stage 1*, which is identified by  $M_{v,q}$ . From this point the expansion is performed in both directions *sequentially* and *independently*. This is because the traffic behavior pattern is generally not “symmetric” in the two directions from the expansion points.

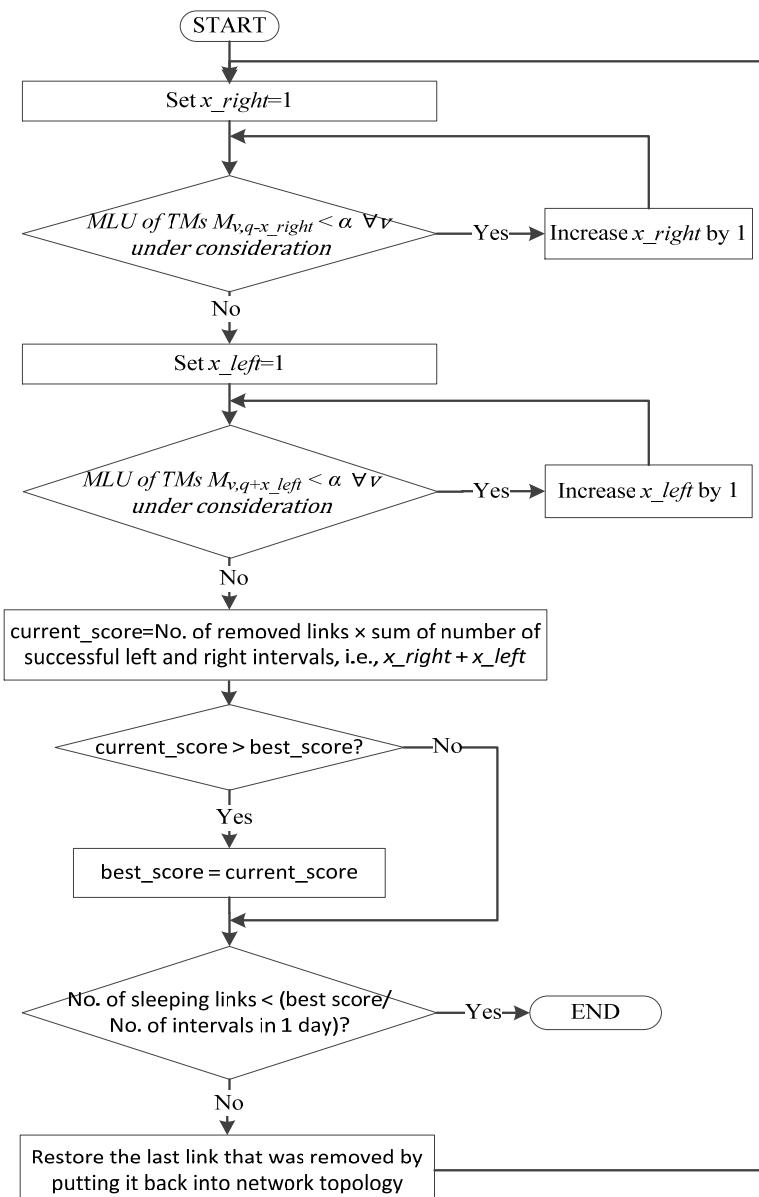


Figure 3-3: Flow chart showing the “off-peak window determination” stage.

In each direction, if the projected MLU of the next TM does not exceed the predefined constraint  $\alpha$  for *all* considered days, then the time duration between that next TM and the current one can be included in the off-peak time window.

After the maximum number of consecutive TMs is obtained, the algorithm reconsiders the possibility of restoring some sleeping links from the reduced topology computed in *Stage 2* and checks if this leads to an improved value of the overall objective function. The consideration of adding sleeping links is in the reverse order by which the links were removed from the network topology in the previous stage. This step is necessary because it is possible to have a sufficient increase in the number of consecutive TMs supported (i.e. enlarged off-peak time window size) to compensate for a reduced number of sleeping links and therefore, obtain better performance according to the objective function.

The termination condition for the iterative consideration of restoring sleeping links depends on the best value of the objective function so far. The algorithm stops if it reaches a point where the number of sleeping links is not sufficient to improve the best value of the objective cost function even if the entire time period is considered off-peak.

For each restoration of a sleeping link, *Stage 3* has to find the new shortest path of all the Source-Destination pairs of the graph and route all the traffic demands of all the traffic matrices until violation of the MLU constraint. Therefore, the overall algorithmic complexity of this stage is  $O(|L| \cdot Q \cdot V \cdot |R|^2 \cdot (|L| + |R| \cdot \log|R|))$ .

## 3.5 Performance Evaluation

### 3.5.1 Experiment Setup

The proposed TLS scheme is evaluated by using the operational GÉANT network topology and its published traffic matrices [4]. Table 3-2 shows the capacity of the links of the GÉANT network topology. The topology is made up of 23 Point-of-Presence (PoPs) and 74 unidirectional links of varying capacity.

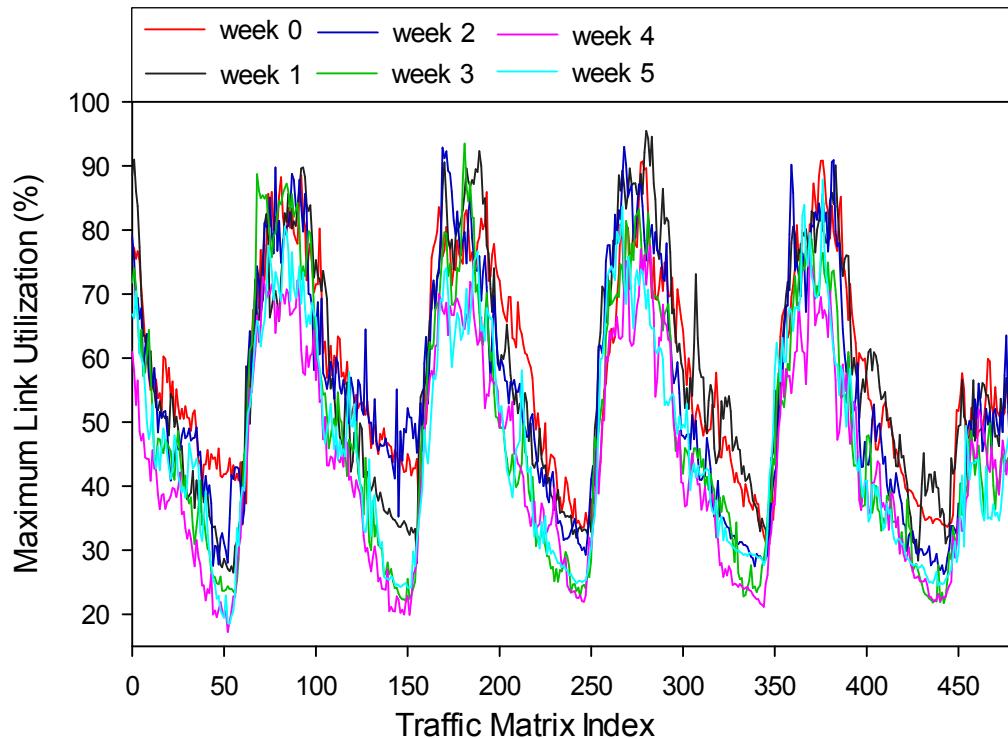
In order to evaluate TLS, the historical traffic matrices of the GÉANT network topology are used [4]. More specifically, six working weeks, with a working week defined in this experiment as the period of time spanning from Monday midday to Saturday midday, are used for the extensive evaluation of TLS. For each working week, 480 consecutive traffic matrices measured at 15-minute intervals are used. The statistical characteristics of the traffic matrices for each week are given in Table 3-3. Week 0 is used for the evaluation of the energy savings obtained by TLS when different MLU constraints,  $\alpha$ , are used.

**Table 3-2: Link characteristic of GÉANT network topology.**

$ L $	$c$ (Mbps)
32	9953
2	4876
32	2488
8	155.2
$\Sigma$	74

**Table 3-3: Characteristics of 6 weeks of the GÉANT traffic matrices.**

MLU (%)	Week					
	0	1	2	3	4	5
Max.	90.9	95.5	93.0	93.5	77.6	87.9
Min.	30.9	26.5	26.3	21.7	17.2	18.5
Mean	58.6	57.0	54.5	48.4	44.0	46.9
1 <sup>st</sup> Quartile	44.5	41.8	40.9	32.7	29.7	31.5
2 <sup>nd</sup> Quartile	55.9	53.9	51.1	44.5	42.2	44.1
3 <sup>rd</sup> Quartile	74.2	73.2	69.7	66.3	57.8	62.0


**Figure 3-4: Maximum Link Utilization versus Traffic Matrix Index for GÉANT.**

The other weeks in the data set are used to verify that TLS can use a representative week, week 0 in this specific scenario, to calculate a reduced topology and its period of operation. Figure 3-4 shows the trend of the MLU of GÉANT over the selected six weeks. The main observation is that

the MLU follows a diurnal pattern where the traffic demands are higher during the day compared to during the night. Furthermore, the troughs and crests of the traffic demands happen about the same time for all the six weeks. These two observations are not unique to GÉANT with several researchers reporting the same observations about other operational networks such as AT&T [5] and Sprint [6].

As previously mentioned, the network operator may have different policies in setting the constraint threshold  $\alpha$  for the maximum allowable link load based on the reduced topology during the off-peak time. This is effectively reflected by the value of  $\alpha$  in Eq. (3.3) in the problem formulation. In the simulation experiments, different values for  $\alpha$ , which represent the conservativeness in the link removal operation, are considered. A higher value of  $\alpha$  indicates that higher MLU is tolerable and in this case, more energy can be saved due to either increased number of links to be removed and/or expanded off-peak time duration allowed. In the experiment, the following four values for  $\alpha$  are considered: 60, 70, 80 and 90, meaning that the maximum allowable MLU is 60%, 70%, 80% and 90% respectively. Certainly the constraint can be further relaxed (i.e. to further increase the value of  $\alpha$ ), however this may lead to increased vulnerability of the reduced network topology, especially due to unexpected traffic upsurge during the off-peak operational period. In addition, the actual traffic demands in the TMs are also scaled up by 10% to 30% in order to evaluate the energy-efficiency with higher traffic volume scenarios. The aim is to make sure that the reduced network topology can also tolerate some variation in MLU values because of traffic uncertainty. The traffic scaling factor is represented by  $\beta$ .

### 3.5.2 Simulation Results

The proposed scheme was evaluated with different constraint thresholds,  $\alpha$ , and traffic scaling factors,  $\beta$ , as shown in Table 3-4. In the table,  $T_o$  is the total off-peak time duration (in terms of minutes) per 24-hour operation and  $E$  represents the energy-efficiency (in terms of percentage) over the whole operation period under consideration.  $E$  is calculated by using the equation (3.5) where  $|A|$  is the number of sleeping links in the “reduced network topology”,  $T_o$  is the time duration during which the “reduced network topology” is operated,  $|L|$  is the total number of links in the network and  $T$  is the total operation time under consideration. It can be observed from Table 3-4 that, as the MLU constraint  $\alpha$  becomes more stringent (i.e.  $\alpha$  becomes smaller), the off-peak window size and number of sleeping links decrease. The window size decreases more significantly compared to the decrease in the number of sleeping links with this specific traffic pattern. This phenomenon is more noticeable when the traffic matrices are scaled up and the MLU constraint is kept constant. The most important result from Table 3-4 is that TLS is able to achieve an energy-efficiency of 28.3% over 5 consecutive working days while having a maximum

MLU constraint of 90% (which is still not higher than the worst-case MLU at peak time that has been observed). In this case, the degree of capacity over-provisioning between peak time and off-peak time is effectively synchronized during daily network operations. In most cases, more conservative MLU threshold settings result in less energy savings being achieved due to the shrunk off-peak window size and/or reduced number of sleeping links. On the other hand, there are observed cases which do not follow this trend. This is because TLS does not always remove the links in the same order in *Stage 2* because the MLU constraint  $\alpha$  is reached at different points in the link removal process because of either a change in the magnitude of the traffic demands due to a change in the traffic scaling factor or a different MLU constraint value. Since every link removal may either reduce or increase the opportunity for further link removals, the energy-efficiency obtained are not always reduced when the traffic scaling factor is increased due to this unpredictability in the future link removal order.

$$E = \frac{|A| \times T_o}{|L| \times T} \quad (3.5)$$

Furthermore, it is also interesting to investigate the change of end-to-end maximum packet delay based on the reduced topologies. In this thesis, the packet delay is measured in terms of the propagation delay since it has been observed that in backbone networks [62][63], the packet delay dominated by the propagation delay and that queuing delay is not significant if the link utilization is below 90% which is the case for the majority of all network scenarios in this thesis. Furthermore, the line-of-sight distance is used for the calculation of the propagation delay. Table 3-5 shows the increase in maximum packet delay in the off-peak window where the reduced network topologies are applied. According to the results, the maximum packet delay experiences only an increase of 0.336% when the MLU constraint  $\alpha$  is set at 90% (without traffic demand scaling). The increase in maximum packet delay under the other MLU constraints,  $\alpha$ , is the same as the 90% scenario because all the Source-Destination pairs in the reduced topologies use the same paths, i.e. the reduced topologies are identical when  $\alpha$  is varied but the off-peak duration changes. The duration of the off-peak windows does not affect the packet delay performance in this scenario.

Figure 3-5, Figure 3-6, Figure 3-7 and Figure 3-8 show the *actual* MLU performance with time across the 5 considered days of week 0 for different  $\alpha$  constraint values. During the peak-time operation period (outside the dark area), the MLU performance is based on the full network topology. There are two MLU curve segments in the off-peak duration: the solid curves represent the MLU performance based on the actual reduced topology with the sleeping links, while the dashed curves indicate the expected MLU performance based on the original full topology.

**Table 3-4: Energy-efficiency of proposed algorithm.**

$\alpha$	$\beta$	$ A_x $	$T_o$	$E(\%)$
90	1	33	915	28.3
	1.1	33	600	18.6
	1.2	33	510	15.8
	1.3	33	420	13.0
80	1	33	600	18.6
	1.1	33	450	13.9
	1.2	33	375	11.6
	1.3	33	360	11.1
70	1	33	435	13.5
	1.1	33	360	11.1
	1.2	33	315	9.76
	1.3	32	405	12.2
60	1	33	330	10.2
	1.1	32	420	12.6
	1.2	32	405	12.2
	1.3	32	375	11.3

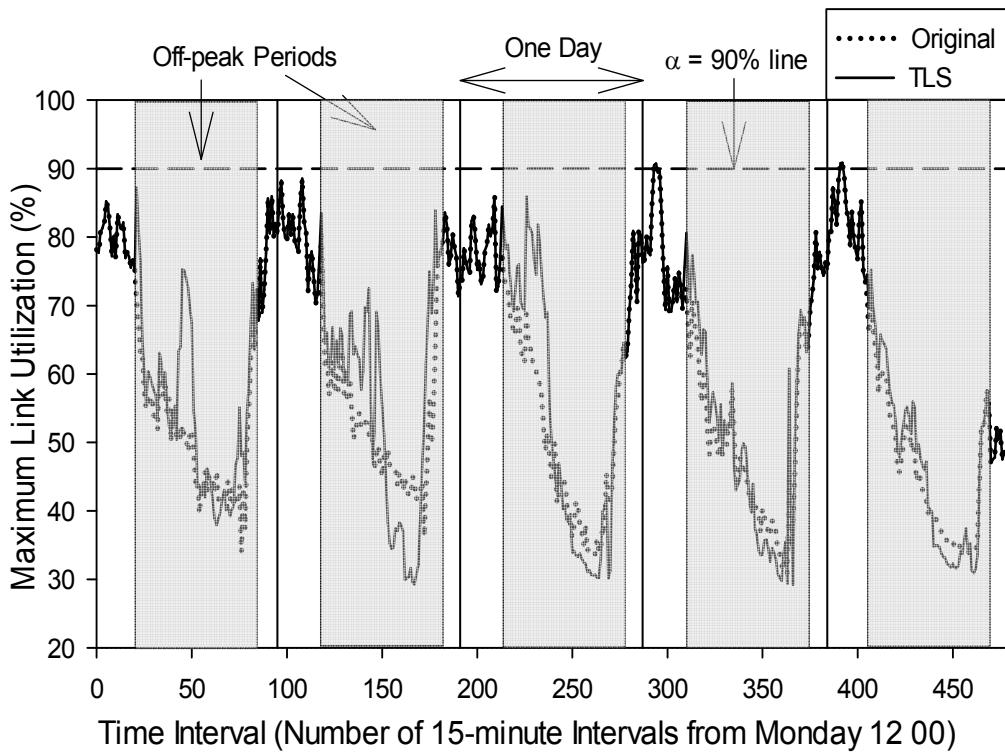
**Table 3-5: Increase in maximum packet delay due to reduced topology.**

$\alpha$	Increase in Maximum Packet Delay in off-peak window (%)
90	0.336
80	0.336
70	0.336
60	0.336

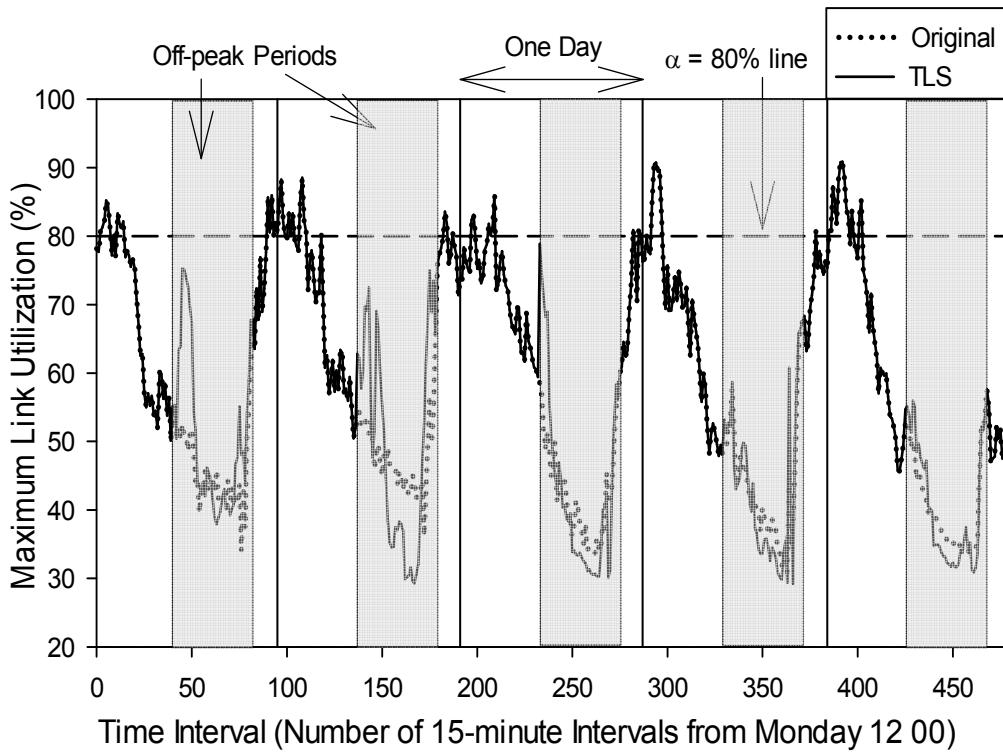
The observation on the MLU comparison is as follows: intuitively, the MLUs in the off-peak windows are expected to be higher when the *reduced network topology* is used rather than the *full network topology*. This is because the number of links available to carry the traffic demands is lower in the former scenario which uses the reduced topology. An interesting observation, however, is that there are also cases where the *full network topology* gives a higher MLU compared to the *reduced network topology*. This is because the greedy link removal operation of TLS may remove a link which originally injects traffic into the most loaded link of the network based on the full topology. The removal of such a link can divert some customer traffic away from that most loaded link to other alternate paths, in which case the overall MLU is decreased.

This scenario is illustrated with the simple illustrative network in Figure 3-9, with its traffic matrix shown in Table 3-6. The figure shows a small topology in which the links have different

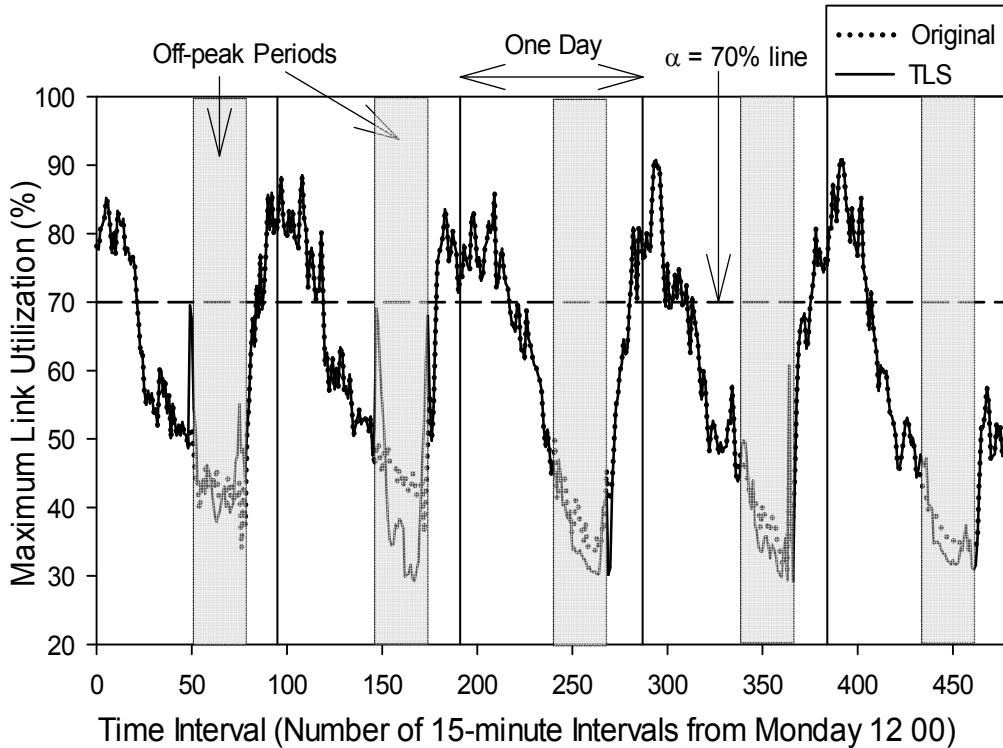
capacities, and the IGP link weights are in *inverse proportion* to the link capacities. Table 3-6 shows the traffic matrix information to be mapped onto the topology. Initially, there is no link removal and the traffic demand from router  $A$  to  $C$  is routed through router  $B$  and it consumes 10 units of bandwidth on the link  $B \rightarrow C$ , which is the most loaded link in the network (note that the link is also carrying 60 units of traffic from  $B$  to  $C$ ) with utilization being 70%. When the least-loaded link  $A \rightarrow B$  is removed from the topology, then the traffic demand from  $A$  to  $C$  will be diverted on to the path  $A \rightarrow D \rightarrow C$ . The link utilization increase on links  $A \rightarrow D$  and  $D \rightarrow C$  due to the diverted traffic (to 30%) is not enough to make any of them become the most loaded link in the network. Although link  $B \rightarrow C$  remains the most loaded link, its utilization is reduced to 60% and hence the maximum MLU in the network is reduced after the link removal.



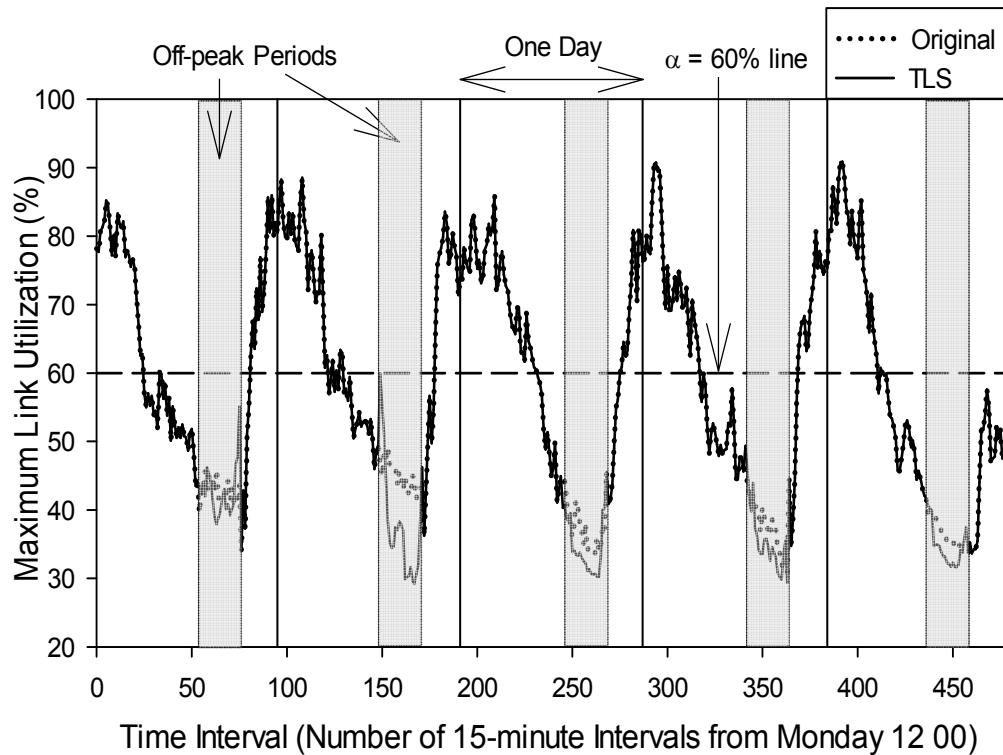
**Figure 3-5: Variation of MLU with reduced topology of  $\alpha=90$ .**



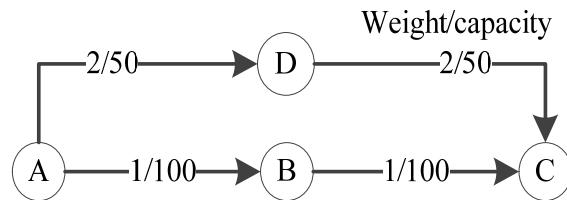
**Figure 3-6: Variation of MLU with reduced topology of  $\alpha=80$ .**



**Figure 3-7: Variation of MLU with reduced topology of  $\alpha=70$ .**



**Figure 3-8: Variation of MLU with reduced topology of  $\alpha=60$ .**



**Figure 3-9: Illustrative network topology to demonstrate the reduction of the MLLU with the reduced topology.**

**Table 3-6: Traffic demands and paths for the illustrative network topology in Figure 3-9.**

Source—Destination	Demand	Initial Path	Path after removal of link A-B
$A - C$	10	$A \rightarrow B \rightarrow C$	$A \rightarrow D \rightarrow C$
$B - C$	60	$B \rightarrow C$	$B \rightarrow C$
$A - D$	20	$A \rightarrow D$	$A \rightarrow D$
$D - C$	20	$D \rightarrow C$	$D \rightarrow C$
<b>MLU</b>	70% ( $B \rightarrow C$ )	60% ( $B \rightarrow C$ )	

Interestingly, it can be also seen in Figure 3-5, Figure 3-6, Figure 3-7 and Figure 3-8 that most of the off-peak MLU values obtained using the original full network topology are similar to the ones obtained using the TLS reduced topology, meaning that the worst-case link load is not increased substantially due to the reduced topology. This is due to the fact that the small amount of load which was previously carried by the sleeping links has been rerouted mostly to the remaining active links with relatively low or medium utilization. This is shown in Figure 3-10 where the utilization percentage of the sorted links is plotted for the original and TLS reduced topology using an off-peak traffic matrix. On the left part of the graph, it can be observed that TLS has lower utilization values due to the sleeping links which have zero utilization while the original topology has a set of links with low utilization. On the right part of the graph though, it can be seen that TLS has larger values than the original due to the loads of the sleeping links being transferred to the other active links.

In addition to the above performance evaluation based on the same 5-day traffic matrices (i.e. week 0) as original input, the configuration obtained for week 0 was applied on 5 other different weeks in order to have a more comprehensive evaluation of the proposed TLS scheme. Table 3-7 shows the performance of the configuration of week 0 when applied to these 5 other weeks where  $H$  is the number of off-peak traffic matrices which produce a MLU above the MLU constraint  $\alpha$  and  $U$  is the actual MLU observed during the off-peak period. It can be observed that the value of  $H$  is always 0 and that  $U$  is always below the constraint  $\alpha$ . These results indicate promising robustness capability of the TLS scheme against varied traffic patterns in different testing weeks.

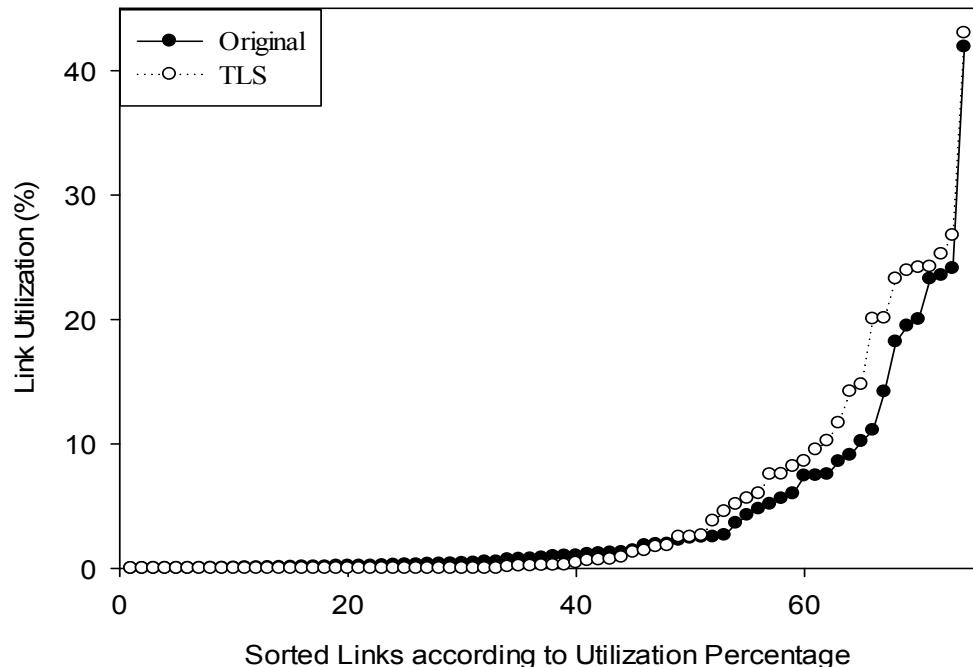


Figure 3-10: Percentage utilization of links in the network.

**Table 3-7: Performance of week 0 configuration on other weeks.**

$\alpha$	Week 1		Week 2		Week 3		Week 4		Week 5	
	$H$	$U(\%)$								
90	0	82.6	0	74.1	0	73.1	0	65.8	0	69.4
80	0	68.5	0	65.5	0	57.5	0	44.5	0	57.5
70	0	59.7	0	60.9	0	50.0	0	42.4	0	48.6
60	0	48.0	0	57.5	0	50.1	0	40.8	0	43.2

### 3.6 TLS with Single Link Failure Protection (TLS-SLFP)

In the proposed TLS scheme presented in first part of this chapter, the reduced network topology is more vulnerable to traffic congestion in the event of any unexpected link failure. These single link failures are known to occur frequently in operational networks [56], and indeed it is important to eliminate, or at least minimize such risk when computing the reduced topology for energy saving purpose. When network devices are put to sleep, the resulting topology has a reduced network capability in carrying customer traffic. In case of unexpected network failures occurring in the residual topology during off-peak time, the operational network may become vulnerable to post-failure traffic congestion. Towards this end, the basic TLS scheme, which was presented in first part of this chapter, is extended to take into account single link failure protections. An advanced scheme called *Time-driven Link Sleeping with Single Link Failure Protection* (TLS-SLFP) is proposed which enables TLS to avoid traffic congestion when any single link failure occurs during the operation of the off-peak reduced network topology. This is achieved through the provisioning of sufficient network resources for carrying ongoing traffic demands without service disruptions when these failures occur. Furthermore, TLS-SLFP ensures that the network still remains always fully-connected in the event of *any* single link failure scenario during the off-peak time.

The second part of this chapter is organized in 4 sections as follows: Section 3.7 provides a brief description of the difference between the problem formulation of TLS and that of TLS-SLFP. In Section 3.8, an overview of the proposed TLS-SLFP scheme is provided along with the necessary alterations and additions to the basic TLS scheme. In Section 3.9, TLS-SLFP is evaluated on the European academic network, GÉANT, and the traffic matrices from week 0 as presented in Section 3.5.1. Furthermore, a comparison between the energy-efficiency performance of TLS and TLS-SLFP is done. Finally, the main findings are summarized in Section 3.10.

### 3.7 Problem Formulation for TLS-SLFP

TLS-SLFP is an extended version of the basic TLS scheme and therefore, both schemes share the same energy-efficiency objective and constraints (the problem formulation of TLS is described in detail in Section 3.2). In contrast to TLS, TLS-SLFP aims to further optimize the reduced network topology so that it always remains fully-connected, and is able to support the worst-case off-peak traffic demands after the failure of *any* single link during off-peak time. First of all, from the basic connectivity point of view, the further reduced (post-failure) network topology should still remain fully-connected. Therefore, an additional constraint must be added to cater for this new requirement. In addition, a still fully-connected post-failure topology does not necessarily ensure that all the traffic demands will be sufficiently accommodated by the further reduced topology. Hence, the second requirement from the viewpoint of capacity support is considered, in which case the original link removal operation needs to make sure that sufficient capacity resources are provided by the reduced topology in the presence of *any* single link failure.

### 3.8 Modifications to the basic TLS Scheme for TLS-SLFP

Since the objective of TLS-SLFP is an extended problem formulation of the plain TLS, the corresponding algorithm is largely built on top of TLS. TLS-SLFP has four stages which are described in detail in the next sections below. TLS-SLFP shares exactly the same *Stage 1* and *3* as TLS and therefore, they are not included here.

#### 3.8.1 Modified Stage 2: Greedy link removal with full network connectivity check for single link failure scenarios

The *Stage 2* of TLS (described in detail in Section 3.4.2) has been extended for taking into account the new constraints on post-failure connectivity of the reduced topology when determining sleeping links. Specifically, when a link  $l^*$  is considered for scheduled sleeping, the following condition needs to be satisfied: The removal of  $l^*$  and of *any* other single link (considered as unexpected single link failure scenarios) should still ensure the residual network topology remains fully connected. As such, the connectivity check is based on  $|L|-2$  single link failure scenarios jointly with the removal of  $l^*$ , where  $|L|$  is the total number of links within the physical network topology.

For each link removal in the modified *Stage 2*, an additional connectivity check needs to be done for each of the remaining active links when they are projected to fail. The algorithmic complexity of the actual connectivity check for any single link failure is  $O(|L|\cdot|R|\cdot(|R|\log|R| + |L|))$ . This

additional step in the modified *Stage 2* results in a new algorithmic complexity of  $O(|L| \cdot (|R|^2 \cdot (|L| + |R| \cdot \log|R|) + |L| \cdot |R| \cdot (|R| \cdot \log|R| + |R|)))$ .

### **3.8.2 New Stage 4: Optimizing sleeping link selection to avoid traffic congestion during single link failures**

A new *Stage 4* has been included in TLS-SLFP which makes sure that no traffic congestion (projected utilization of any active link exceeding 100%) occurs given any single link failure scenario in the robust reduced topology for the off-peak traffic demands. All the steps in *Stage 4* are shown in Figure 3-11. The first step of this stage calculates the number of single link failure scenarios that *cannot* be supported by the reduced network topology produced at the end of *Stage 3*. In order to test for single link failures in this reduced network topology, each active link is removed in turn from the topology and the ability of each residual topology to accommodate the traffic demands is verified.

After identifying the total number of unsupported single link failure scenarios, TLS-SLFP identifies what impact each sleeping link in the reduced topology will have on this number if the sleeping link were to be restored back to the topology again. This step produces a list of ranked sleeping links for potential restoration according to their ability to reduce the number of unsupported single link failure scenarios. The next step of this stage involves the restoration of sleeping link candidates back to the working topology according to the list produced in the previous step. After each sleeping link candidate is considered for restoration, the number of unsupported single link failure scenarios is re-calculated, and if this number has not decreased, the link can remain sleeping and the next link candidate in the list is tested.

This procedure is applied to ensure that the minimum number of sleeping link candidates needs to be restored back to the working topology, and therefore, conserve the energy-efficiency achieved in *Stage 3* as much as possible.

In this stage, the first step involves the calculation of the number of single link failures which cause the MLU to go above 100% in order to see if optimization of the number of sleeping links is needed. This step has algorithmic complexity of  $O(|L| \cdot Q.V. |R|^2 \cdot (|L| + |R| \cdot \log|R|))$ . The next step involves the calculation of the number of unsupported single link failures for each sleeping link restoration. The algorithmic complexity of this step is  $O(|L|^2 \cdot Q.V. |R|^2 \cdot (|L| + |R| \cdot \log|R|))$  which dominates the overall algorithmic complexity of this stage. The next major step of *Stage 4* is to calculate the remaining unsupported single link failures after each sleeping link restoration. The algorithmic complexity of this step is the same as the previous step. Therefore, the overall algorithmic complexity of this stage can be simplified to  $O(|L|^2 \cdot Q.V. |R|^2 \cdot (|L| + |R| \cdot \log|R|))$ .

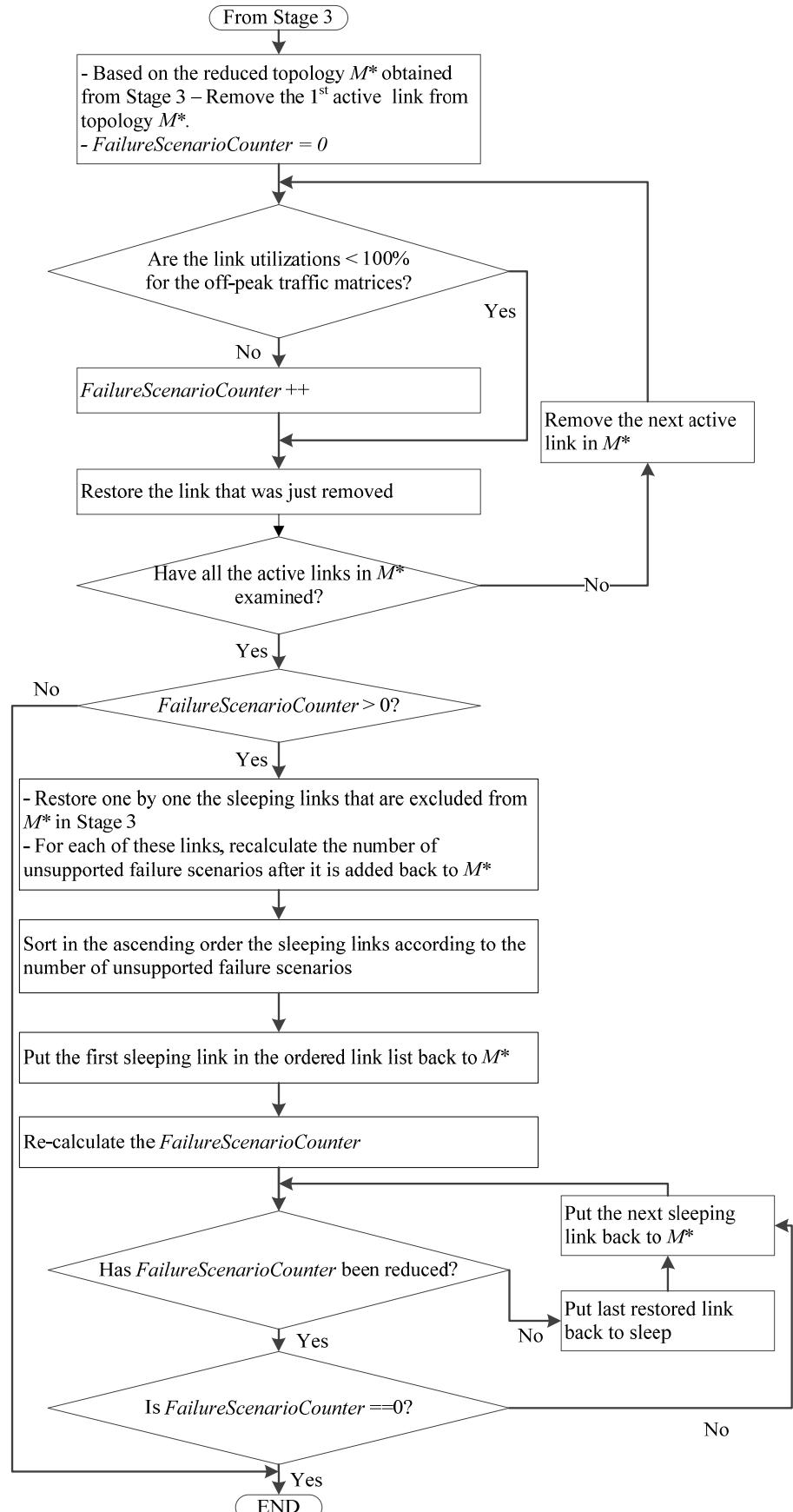


Figure 3-11: Flow chart showing optimization of the set of sleeping links in TLS-SLFP.

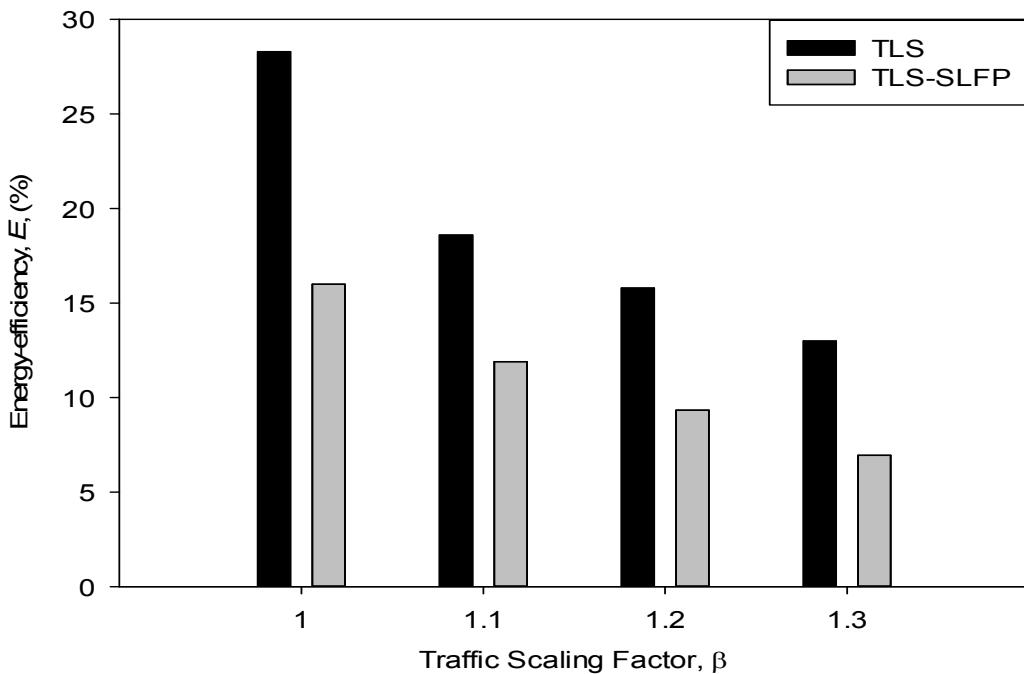
### 3.9 Performance Evaluation

The TLS-SLFP scheme was evaluated based on the same GÉANT network topology and traffic matrices, from week 0, that were used for TLS (described in detail in Section 3.5.1), since this allows for the change in performance to be directly compared. Figure 3-12, Figure 3-13,

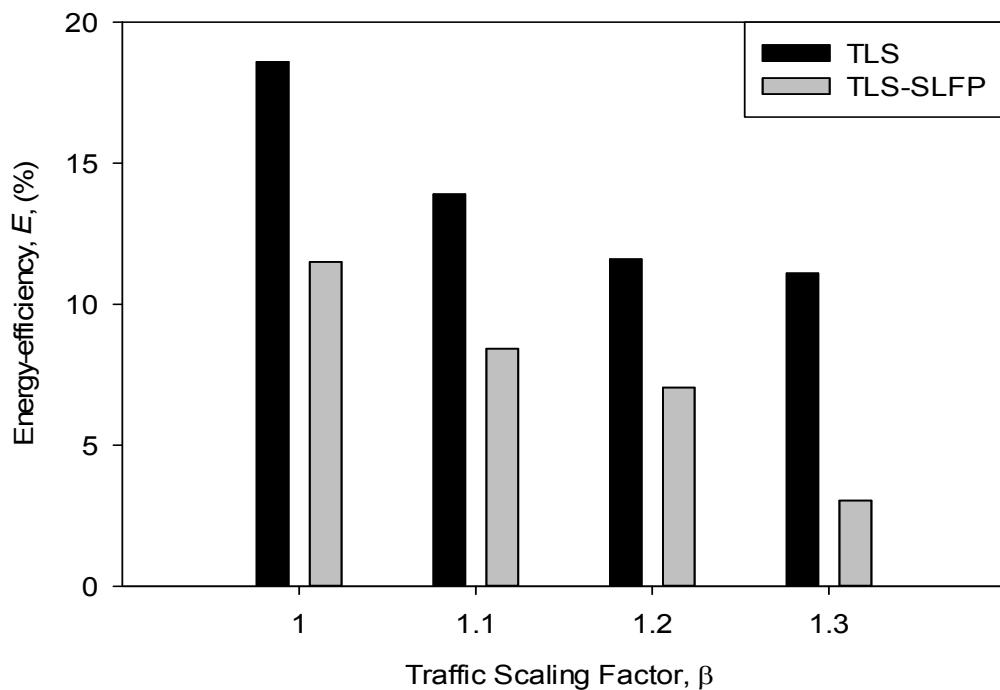
Figure 3-14 and Figure 3-15 indicate the energy saving comparison between TLS and TLS-SLFP based on different values of MLU constraint  $\alpha$ . All the figures indicate that the overall energy saving gains are reduced as new constraints are added to TLS in order to take into account single link failure protections in TLS-SLFP.

Figure 3-12 shows the performance comparison between TLS and TLS-SLFP when the MLU constraint is set to 90%. It can be seen that in all cases involving different traffic scaling factors, there is a decrease in energy savings when the additional constraint is applied for single link failure protection. For instance, the energy saving gain is reduced from 28.3% in TLS to 16% in TLS-SLFP in the scenario without traffic volume scaling. Figure 3-13,

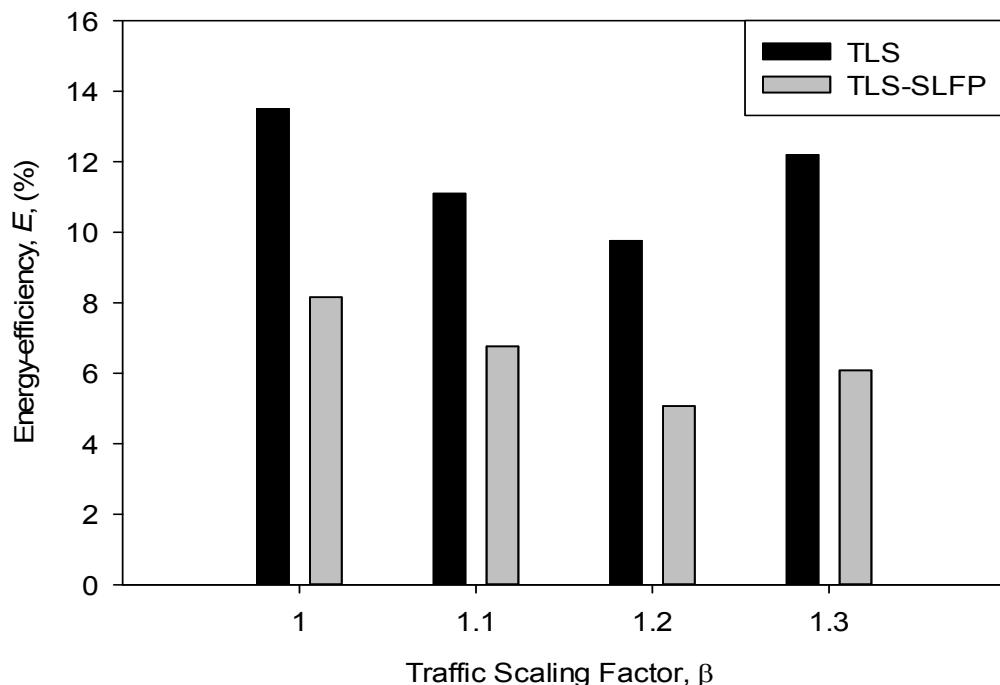
Figure 3-14 and Figure 3-15 where the MLU constraint is set to 80%, 70% and 60% respectively, show a similar pattern as observed in Figure 3-12.



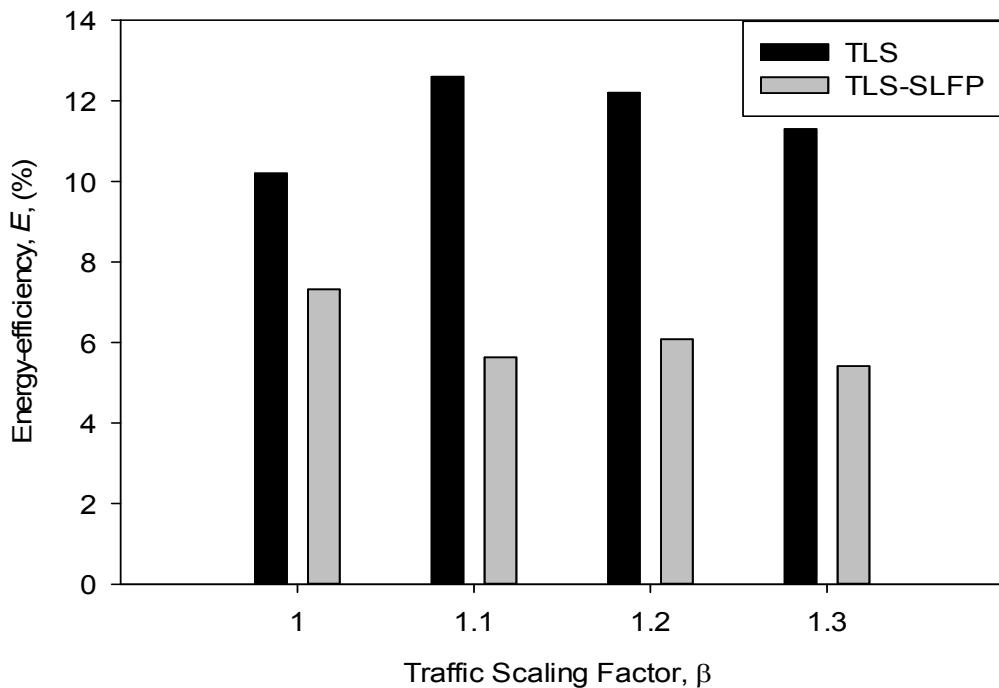
**Figure 3-12: Energy saving gains with reduced topology of  $\alpha=90$  under different traffic scaling factors.**



**Figure 3-13.** Energy saving gains with reduced topology of  $\alpha=80$  under different traffic scaling factors.



**Figure 3-14.** Energy saving gains with reduced topology of  $\alpha=70$  under different traffic scaling factors.



**Figure 3-15. Energy saving gains with reduced topology of  $\alpha=60$  under different traffic scaling factors.**

**Table 3-8: Energy-efficiency of TLS-SLFP.**

$\alpha$	$\beta$	$ A_x $	$T_o$	$E(\%)$
90	1	12	1425	16.0
	1.1	18	705	11.9
	1.2	17	585	9.33
	1.3	17	435	6.94
80	1	13	945	11.5
	1.1	13	690	8.42
	1.2	20	375	7.04
	1.3	9	360	3.04
70	1	20	435	8.16
	1.1	20	360	6.76
	1.2	20	270	5.07
	1.3	16	405	6.08
60	1	20	390	7.32
	1.1	20	300	5.63
	1.2	16	405	6.08
	1.3	16	360	5.41

Table 3-8 shows a detailed breakdown of the performance of TLS-SLFP. One interesting observation from this table is that TLS-SLFP tends to push more links to the sleep mode when the MLU constraint is lowered while decreasing the off-peak window duration. Overall, the energy savings decrease as the MLU constraint is lowered because the MLU threshold is reached more quickly during the off-peak window expansion process. It is also worth mentioning that the energy saving gains does not always decrease when the MLU constraint is kept constant and the scaling factor is increased. This is due to the same reason in the TLS scenario where the link removal order changes as MLU constraint and scaling factor change.

### **3.10 Implementation by Network Operators**

Network operators can verify whether TLS, or its failure-resilient version TLS-SLFP, is suitable for their network by first measuring the traffic pattern in their network. As mentioned previously, TLS is designed for backbone networks which have a regular and predictable traffic pattern and if network operators observe that their network behaves in this manner, then TLS can be a good candidate ETE scheme for their network. In most cases, network operators do not need to install any new monitoring tools in their network since they already have the traffic pattern of their network for network planning purposes. After verifying that the traffic pattern of their network is diurnal and regular, network operators can input this information along with information about the network topology of their network into a network simulator which is capable of running the TLS/TLS-SLFP scheme to get an indicative energy-savings, maximum link utilization and maximum packet delay. If the network operators are satisfied with the results obtained, they can go forward and implement TLS/TLS-SLFP in their network.

TLS is flexible enough ETE scheme that it allows network operators to make a trade-off between the amounts of energy-savings that they want to achieve and the maximum link utilization that they are willing to accept as a consequence of the greater energy-savings obtained. The trade-off that network operators are willing to accept will depend on their network policy. The network policy of the network operator will also determine whether TLS-SLFP should be implemented instead of plain TLS because of the capability of TLS-SLFP to protect against packet loss during single link failures. While TLS-SLFP offers greater protection against single link failures compared to TLS, the amount of energy-savings is less.

### **3.11 Summary**

In this chapter, the motivation behind developing a simple time-driven energy saving scheme for backbone networks is presented. The proposed heuristic algorithm, TLS, determines a *reduced network topology* with scheduled sleeping link re-configuration and the actual off-peak time

period during which this network topology is used. The scheme aims to identify an optimized trade-off between the number of sleeping links and the configuration duration for the reduced topology in order to achieve maximum energy savings performance. It is demonstrated through simulations based on the GÉANT network and its real traffic matrices, that the proposed scheme is able to achieve significant energy savings in the daily operation. Furthermore, there is low impact on the end-to-end maximum packet delay when applying the reduced network topology during off-peak time. Some interesting observations have also been obtained and analyzed, such as the ability of the proposed algorithm to reduce the MLU with fewer active links, which provide insights for the design of future traffic-engineering aware network optimization schemes for energy-efficiency.

In the second part of this chapter, a new algorithm called TLS-SLFP was introduced as an extension to TLS to take into account single link failure protection which is known to occur frequently in operational networks. TLS-SLFP optimizes the off-peak reduced network topology with the main objectives of avoiding the network from becoming disconnected and congested due to unexpected single link failures. Even though there have been additional constraints on the scheme such as more stringent network connectivity and spare capacity in order to support single link failure events, TLS-SLFP has been shown to be able to achieve reasonable energy savings.

## Chapter 4

# 4 Optimizing IGP Link Weights for Energy-efficiency and Load-balancing in ISP Networks<sup>3</sup>

## 4.1 Introduction

In addition to energy-efficiency, load-balancing has been a common objective of plain traffic engineering in computer networks, and numerous schemes have been developed towards this objective [64]. Load-balancing aims to reduce the Maximum Link Utilization (MLU) in the network through an optimized distribution of traffic. This reduction in MLU allows the networks to offer better Quality of Service (QoS) assurance and also to efficiently handle unexpected upsurges in traffic demands. However, since load-balancing attempts to “spread” the traffic while ETE algorithms attempt to “concentrate” traffic only to residual active devices (e.g. network links), conventional load-balancing and ETE are intuitively conflicting each other in network configurations.

In this chapter, a new ETE scheme, called *Green Load-balancing Algorithm* (GLA), is proposed which jointly optimizes the load-balancing and energy-efficiency in the network based on existing ETE schemes. GLA achieves such objectives by optimizing the Interior Gateway Protocol (IGP) link weights of a network, which influences the distribution of traffic in a network. Existing ETE schemes do not re-optimize the IGP link weights after they put some links to sleep and this is not efficient because the IGP link weights are only optimal when all the links in the network are active, i.e. not sleeping. Specifically, the intelligent setting of IGP link weights in GLA is able to *maximize energy saving gains* through link sleeping, while *maintaining, or even further*

---

<sup>3</sup> The content of this chapter is based on the following publications:

1. F. Francois, N. Wang, K. Moessner, S. Georgoulas, and K. Xu, "Green IGP Link Weights for Energy-Efficiency and Load-balancing in IP Backbone Networks", in Proc. of 2013 of IEEE/IFIP Networking, May 2013.
2. Frederic Francois, Ning Wang, Klaus Moessner, Stylianos Georgoulas and Ke Xu, "On IGP Link Weight Optimization towards joint Energy-efficiency and Load-balancing," Elsevier Computer Communications (**under review**).

*improving load-balancing performance* on top of the residual working topology. Hence, GBP is different from conventional single-objective IGP link weight setting schemes which optimize only for load-balancing but do not efficiently provide opportunities for link sleeping operations. In the second section of this chapter, the end-to-end Maximum Packet Delay (MPD) performance of three existing ETE schemes are investigated and found to be unsuitable for the strict QoS requirements that network operators have. In the light of this observation, the three existing ETE schemes are modified so that they can respect an upper bound on MPD. The new MPD-aware ETE schemes are then evaluated and compared against their original counterparts in terms of load-balancing and energy savings performance.

In order to illustrate the basic concept of GLA-based link weight optimization for both load-balancing and energy-efficiency, the small illustrative network topology in Figure 4-1, with indicated link capacities and IGP weight settings, is used. The aim of such an example is to illustrate how IGP link weights can be manipulated in order to create opportunities for more links to sleep, but without affecting the load-balancing requirements. For simplicity and clarity, an “incomplete” uni-directional graph is used, but certainly such an idea is also applicable to real network topologies with full bi-directional connectivity. First of all, it can be observed that there are only two links which can be put to sleep without causing the network topology to lose full connectivity: namely links  $C \rightarrow D$  and  $A \rightarrow D$ . First, the case where the set of link weights is non-optimized as shown on the left is considered, and then the case where the link weights are optimized on the right side of the figure is considered.

Given a simple illustrative traffic matrix composed of only traffic demands between two Source-Destination pairs:  $C - D$  and  $A - D$  of 30 and 75 units respectively, the traffic demand  $C - D$  goes through path  $C \rightarrow D$  with a link utilization of 60%. For the traffic demand  $A - D$ , path  $A \rightarrow D$  is used with a link utilization of 75%. The original MLU in this scenario is therefore 75%. Based on the given traffic demands, the conventional techniques such as [11] are followed to consider link removal one by one from the topology where the least utilized link is selected first. First of all, if link  $C \rightarrow D$  is put to sleep, its load is re-routed through the alternative path  $C \rightarrow A \rightarrow D$  and the resulting utilization on link  $A \rightarrow D$  becomes 105%. Hence,  $C \rightarrow D$  cannot be put to sleep because it causes the network to become overloaded. Link  $A \rightarrow D$  can be put to sleep because the alternative path  $A \rightarrow B \rightarrow D$  will have MLU of 37.5%. Therefore, the resulting MLU in the network is 60% and only one link can be put to sleep.

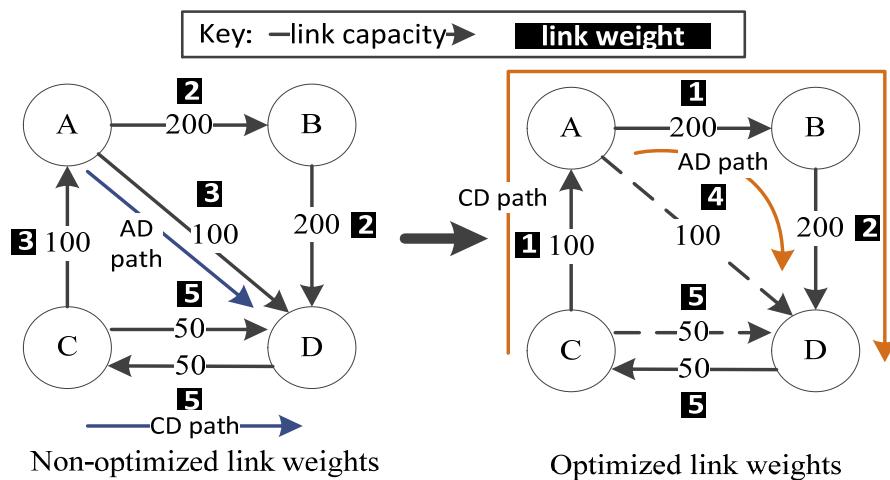
If the link weights are optimized as shown on the right side of Figure 4-1, both links  $C \rightarrow D$  and  $A \rightarrow D$  can be put to sleep without causing the network to become overloaded. The traffic demands  $C - D$  and  $A - D$  are routed along the new paths  $C \rightarrow A \rightarrow B \rightarrow D$  and  $A \rightarrow B \rightarrow D$  respectively. The MLU of the resulting network becomes 52.5% even though two links are now

put to sleep. The optimization of the link weights has made it possible not only to reduce the MLU from 75% to 52.5%, but also allows one more link to go to sleep, achieving simultaneous improvement of both load-balancing and energy-efficiency objectives.

In addition to the aforementioned simple example that illustrates the effectiveness of GLA based on one snapshot of traffic matrix considered in some simple ETE schemes [17][65], the proposed scheme can be further applied to more advanced ETE schemes which take into account traffic dynamics. According to the *Time-driven Link Sleeping* (TLS) scheme proposed in Chapter 3 [66][67], both the set of sleeping links and their sleeping time are jointly determined by taking into account the patterns of traffic demands for a given period of time.

When GLA is to be applied to such schemes, additional constraints need to be considered. Specifically, a *common* set of IGP link weights needs to be applied across a *diverse* set of traffic matrices. This requires the algorithm to be robust enough to ensure optimized load-balancing performance is achieved across different time periods.

The rest of this chapter is organized as follows: in Section 4.2, the *generic* problem of the joint optimization of load-balancing and energy-efficiency is formulated. A brief review is given on the working mechanism of three different ETE schemes based on which GLA will be evaluated. In Section 4.3, the detailed specification of the GLA algorithm is presented. In Section 4.4, GLA is evaluated by using the Point-of-Presence representation of the European academic network GÉANT and its real traffic matrices. In Section 4.5, an enhanced version of GLA, which can be more natively coupled with one of the considered ETE schemes: *Time-driven Link Sleeping*, is introduced. In Section 4.6, the Maximum Packet Delay (MPD) performance of the different ETE schemes is investigated. In Section 4.7, the existing ETE schemes are modified so as to improve their MPD performance and Section 4.8 describes the change in the load-balancing and energy-efficiency performance due to these modifications. Finally, a summary of the main findings of this chapter is provided in Section 4.9.



**Figure 4-1: Illustrative network topology to illustrate optimization of link weights.**

## 4.2 Problem Formulation and Existing ETE Schemes

### 4.2.1 Problem Formulation

**Table 4-1: Definition of symbols for GLA.**

Variable	Description
$\alpha$	Maximum allowable utilization of link capacity
$c_{ij}$	Bandwidth capacity of link from router $i$ to $j$
$E$	Energy-efficiency. Each ETE scheme has its own definition of energy-efficiency which is described in more detail in Section 4.2.2.
$f_{ij}^{sd}$	Traffic demand from router $s$ to $d$ that traverses link from router $i$ to $j$
$f_{ij}$	Total traffic demand on link from router $i$ to $j$
$G(R, L)$	Directed graph with $R$ being set of routers and $L$ being set of links
$t^{sd}$	Traffic demand from router $s$ to $d$
$U$	Maximum link utilization

The joint optimization of load-balancing and energy-efficiency in a network can be expressed with the following two objectives:

$$\text{minimize } U \quad (4.1)$$

$$\text{maximize } E \quad (4.2)$$

subject to:

$$\sum_{j=1}^{|R|} f_{ij}^{sd} - \sum_{j=1}^{|R|} f_{ji}^{sd} = \begin{cases} t^{sd} & \forall s, d, i = s \\ -t^{sd} & \forall s, d, i = d \\ 0 & \forall s, d, i \neq s, d \end{cases} \quad (4.3)$$

$$f_{ij} < \frac{\alpha}{100} \times c_{ij} \quad \forall i, j \text{ with } \alpha \in [0, 100] \quad (4.4)$$

Equation (4.1) represents the first objective of GLA which is the minimization of the Maximum Link Utilization (MLU) in the network in order to achieve load-balancing. For the GLA scenarios based on one traffic matrix, it refers to the MLU related to that snapshot only. For the scenarios involving multiple traffic matrices (e.g. based on TLS), this refers to the *worst-case scenario* across all considered traffic matrices. Equation (4.2) represents the second objective of GLA which is the maximization of the energy-efficiency given by an existing ETE scheme. Again, the definition of  $E$  is specific to the individual ETE schemes considered, which will be introduced in Section 4.2.2. Equation (4.3) represents the standard flow conservation constraint. Equation (4.4) ensures that whenever a reduced topology is used, all active links should have their utilization below a given threshold  $\alpha$  as determined by the ISP. That is, with a set of links being put into

sleep mode, the maximum link load should not exceed the threshold  $\alpha$  (in terms of a percentage of the link capacity). Another constraint is that the network needs to remain fully connected when links are configured to sleep mode so that there is always a path between any two routers in the reduced network topology.

### 4.2.2 Existing ETE Schemes

In this section, a brief review of the three different offline ETE schemes, based on which GLA is applied, is given. The first two schemes, Least Flow (LF) and Most Power (MP), were introduced in [17] and [65] respectively and were chosen to evaluate GLA because of their simplicity. The third ETE scheme over which GLA was evaluated is *Time-driven Link Sleeping* (TLS), which is the ETE scheme introduced in Chapter 3. The major difference between the first two schemes and TLS is that the first two schemes operate on one single traffic matrix snapshot at a time and do not consider a collection of dynamic traffic matrices as TLS does. It is also worth mentioning that each of these three schemes has a different way of calculating the Eq. (4.1) and (4.2) in Section 4.2.1.

#### 4.2.2.1 The Least Flow Scheme

In [17], the authors use an ETE scheme called Least Flow (LF). LF iteratively selects the least loaded link in the network as candidate for sleeping. The selected link can only go to sleep if the full connectivity of the network topology is maintained and the resulting MLU is below a given threshold  $\alpha$  when the link enters sleep mode. Otherwise, the next least loaded link is selected for sleeping consideration until all the links have been investigated.

During the operation of GLA on top of LF, the value of Eq. (4.1) is the MLU value when the single traffic matrix is mapped onto the network with its full topology (i.e. all the links in the network are active). The MLU value obtained is then used as the value of  $\alpha$  in Eq. (4.4). The energy-efficiency (value of Eq. (4.2)) of LF is calculated as shown by Eq. (4.5) where  $|A|$  is the number of sleeping links and  $|L|$  is the total number of links in the network. The algorithmic complexity of the LF scheme was given as  $O(|L|. |R|^2. (|L| + |R|. \log|R|))$  in [17].

$$E = \frac{|A|}{|L|} \quad (4.5)$$

#### 4.2.2.2 The Most Power Scheme

The Most Power (MP) ETE scheme presented in [65] is similar to LF. MP iteratively selects the link which consumes the highest amount of power in the network as candidate for sleeping. The selected link can only go to sleep if the full connectivity of the network is maintained and the

resulting MLU is below a given threshold  $\alpha$  when the link is sleeping. Otherwise, the next link which consumes the most power becomes candidate for removal.

During the operation of GLA on top of MF, the value of Eq. (4.1) is calculated in the same way as for the LF algorithm above. The MLU value obtained is then used as the value of  $\alpha$  in Eq. (4.4). The energy-efficiency (value of Eq. (4.2)) of MP is calculated as shown in Eq. (4.6) where  $A$  is the set of sleeping links,  $P_l$  is the power consumed when the link with index  $l$  is active and  $B$  is the set of links that are sleeping. As mentioned previously, the energy savings due to a link being put into sleep mode can pre-dominantly be attributed to the line cards which are connected to the link. The algorithmic complexity of the MP scheme was given as  $O(|L| \cdot |R|^2 \cdot (|L| + |R| \cdot \log|R|))$  in [65].

$$E = \frac{\sum_{l=0: l \in A}^{|L|} P_l}{\sum_{l=o}^{|L|} P_l} \quad (4.6)$$

#### 4.2.2.3 The Time-driven Link Sleeping Scheme

The third existing ETE scheme over which GLA is evaluated upon is the TLS scheme described in detail in Chapter 3. In the operation of GLA on top of TLS, the objective value,  $U$ , in Eq. (4.1) is equal to the worst-case MLU in the network when all traffic matrices are considered. Specifically, the metric  $U$  in Eq. (4.1) represents the *peak-time* MLU in the network to be optimized. The MLU constraint for the off-peak time, represented by  $\alpha$  in Eq. (4.4), depends on either the obtained peak-time MLU or is pre-determined by the network operator. The energy-efficiency (value of Eq. (4.2)) of TLS is calculated according to Eq. (4.7), where  $|A|$  is the number of sleeping links in the “reduced network topology”,  $T_o$  is the time duration during which the “reduced network topology” is operated,  $|L|$  is the total number of links in the network and  $T$  is the total operation time under consideration. According to Eq. (4.7), the energy-efficiency of TLS can only be increased by increasing the nominator in Eq. (4.7) since the denominator is fixed. Intuitively, an increase in the number of sleeping links may lead to a smaller off-peak duration (i.e. sleeping time  $T_{op}$ ), because the capacity of network is reduced and only a smaller number of traffic demands can now be satisfied. Therefore, a trade-off needs to be obtained between  $|A|$  and  $T_o$ . The algorithmic complexity of the TLS scheme was given as  $O(|L| \cdot Q \cdot V \cdot |R|^2 \cdot (|L| + |R| \cdot \log|R|))$  in Chapter 3 where the symbols  $Q$  and  $V$  are the number of traffic matrices collected in each day and the number of days considered respectively.

$$E = \frac{|A| \times T_o}{|L| \times T} \quad (4.7)$$

## 4.3 Green Load-balancing Algorithm (GLA)

### 4.3.1 Scheme Overview

It is well-known that computing the optimal link weights for basic load-balancing is already an NP-hard problem [52] and therefore, a novel ETE scheme called *Green Load-balancing Algorithm* (GLA) is proposed. GLA uses meta-heuristics (evolutionary/genetic algorithms) to find the optimized IGP link weights which can solve the more complicated problem of the joint-optimization of load-balancing and energy-efficiency in a backbone network.

GLA is used to solve the problem of finding the set of optimized green IGP link weights which caters for both objectives 1 and 2 as represented by Eq. (4.1) and (4.2) respectively in Section 4.2.1. GLA is implemented in the form of a customized version of the Non-dominant Sorting Genetic Algorithm (NSGA-II) [68]. NSGA-II operates in a similar fashion to traditional genetic algorithms. NSGA-II has been chosen because it is a multi-objective algorithm which preserves diversity and elitism of the solution space and has low algorithmic complexity. NSGA-II can be used to find the Pareto-optimal front of a solution space and therefore, there is no need to use a single aggregated objective function as is required with other meta-heuristics which can accommodate only one single objective function. Moreover, the ability to do multi-objective optimization allows network operators to postpone their decision on their desired trade-off between load-balancing and energy-efficiency until the end of the GLA scheme because the GLA schemes will provide as final result the different Pareto-optimal points that the network operator can choose. The Pareto-optimal front arises due to the presence of two objectives in the problem formulation, load-balancing and energy-efficiency. Intuitively, load-balancing and energy-efficiency through link sleeping are two conflicting objectives. This is because load-balancing aims to reduce the load on highly-utilized links by shifting the traffic demands on these links to less utilized links in the network while energy-efficiency requires the traffic to be concentrated on a subset of active links and putting the non-utilized links to sleep, which results in the active links becoming highly-utilized. The conflict between the two objectives gives rise to a Pareto-optimal front in the sense that, when the MLU is reduced for load-balancing purposes, the energy-efficiency objective will be sacrificed because of the more constrained environment for the different ETE schemes. Hence, an optimized trade-off needs to be obtained between these two objectives. In addition to the basic NSGA-II operations, two custom operators are also introduced to further enhance the performance of GLA by enabling a more efficient search to be done in the solution space.

### **4.3.2 Solution Encoding**

In the genetic algorithm, the solution (i.e. the set of IGP link weights in GLA) is encoded through a chromosome. A chromosome is made up of a number of genes, which is equal to the number of links in the network in GLA. Therefore, each gene in the chromosome represents a link in the network. Each gene is restricted to an integer value in the range of 1 to 65535. This range corresponds to the range of values allowed for IGP link weights. In GLA, a link with a link weight of 65535 is defined as sleeping and is not used to route traffic demands.

### **4.3.3 Fitness Functions**

Each chromosome (i.e. solution candidate) has two distinct fitness functions in the NSGA-II algorithm which are represented by Eq. (4.1) and (4.2) in Section 4.2.1 respectively. As mentioned previously, since the three ETE schemes, described in Section 4.2.2, have different mechanisms, their fitness functions differ in the way they calculate Eq. (4.1) and (4.2) (described in Section 4.2.1).

### **4.3.4 Sleeping Link Crossover Operator**

A crossover operation in a genetic algorithm involves taking two chromosomes (i.e. two solutions) in the current population (set of solutions) and swapping their genes (i.e. link weights) with each other to produce two new offspring chromosomes. The aim is to produce newly generated solution candidates with better fitness values. In order to explore the solution space more efficiently and achieve quicker convergence, a customized crossover operator is designed in addition to the standard operators such as two-segment crossover. This new crossover operator has been designed so that one parent chromosome can replace its gene with its counterpart gene in the other parent chromosome if the counterpart gene represents a sleeping link.

Figure 4-2 shows the operation of the new crossover operator. A chromosome in this operation is made up of two rows. The top row contains the IGP link weights while the bottom one contains a binary array which indicates link status. In this array, a “1” value means that the associated link is sleeping and “0” means the associated link is active. The IGP weight of an active link in a chromosome is changed with its counterpart in the other parent chromosome only if its counterpart is marked as sleeping. The arrows between the two parent chromosomes in Figure 4-2 show when link weight change occurs and the direction of the change. This operation is similar to “XORing” the bottom row of the two parent chromosomes. This new crossover operator allows “good” link weights which promote link sleeping to propagate through the population.

### 4.3.5 Link Utilization Mutation Operator

A mutation operation in a genetic algorithm involves taking one chromosome (i.e. solution candidate) in the population and modifying one or more genes (i.e. link weights) in the chromosomes. This operation is done so that new genes, which do not exist in the current population, are introduced in the population with the aim that this will likely increase the fitness functions of the selected chromosomes. A new mutation operator has been developed based on the percentage utilization of each link when a traffic matrix is mapped onto the topology. In this new mutation operator, each link weight has  $Pr_{g\_mut}$  probability of mutating which is equal to the utilization percentage of that link divided by the MLU in the network.

As such, highly utilized links will have a higher probability of mutating compared to links with low utilization. The probability that a chromosome in the parent population will begin this utilization-based mutation is given by  $Pr_{mut}/2$  where  $Pr_{mut}$  is the probability of a chromosome undergoing mutation and there is  $Pr_{mut}/2$  probability that the chromosome will undergo the other standard mutation operators.

If a link is selected to be mutated, its link weight will be increased according to Eq. (4.8) where  $W_n$  is the new link weight.  $W_o$  is the old link weight and is used as the mean of a normal distribution  $N$  with standard deviation  $\varepsilon W_o$  where  $\varepsilon$  is a fractional multiplier lower than 1. The rationale behind Eq. (4.8) is that highly utilized links will have their link weights increased, and hence make them less likely to be chosen by IGP for routing traffic demands. This is likely to cause the load of these links to decrease if alternative shorter paths are identified to route the traffic demands.

$$W_n = W_o + |W_o - N(W_o, \varepsilon W_o)| \quad (4.8)$$

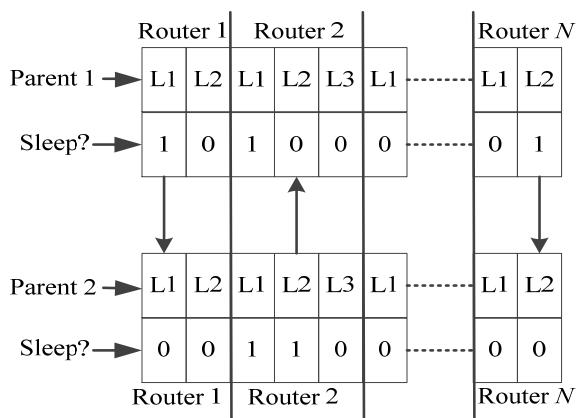


Figure 4-2: Sleeping link crossover operator.

### **4.3.6 Overall Operation of GLA**

At the beginning of GLA, an initial population of random chromosomes is generated. An offspring population is then created through the joint application of crossover and mutation operators on some randomly selected chromosomes in the parent population. In addition to the traditional crossover and mutation operators such as two-segment crossover and random gene mutation, the two custom crossover and mutation operators, described in the previous two subsections, are applied. These customized genetic operators help the exploration of the search space in a more efficient manner. The parent and offspring populations are then merged and sorted according to the fitness and diversity scores of the individual chromosomes to create a new parent population. The algorithm stops when a given targeted number of generations has been calculated, or there has been no improvement in the obtained solutions since a given number of generations. The problem-independent complexity of NSGA-II is reported to be  $O(\kappa\chi^2)$  [68] where  $\kappa$  is the number of objectives in the problem (i.e. two for GLA) and  $\chi$  is the population size. The complexity of the evaluation of the fitness functions for each solution in the population depends on the ETE schemes used and has been reported in Section 4.2.2.

## **4.4 Performance Evaluation**

### **4.4.1 Network Scenario**

The performance of GLA when applied to the different ETE schemes ([57][69][66][67]) is evaluated by using the operational network topology GÉANT and its published traffic matrices [4]. The published topology consists of 23 Points-of-Presence (PoPs) and 74 unidirectional links of varying bandwidth capacities which are described in Table 4-3 below. The total power consumption due to a link being put into sleep mode,  $p$ , is also given in the table.

These values were calculated from the power consumption model of line cards in [70] with the assumption that line cards are responsible for most of the power consumption of a link [10][11]. The power values are used by the 2<sup>nd</sup> ETE algorithm, MP, to decide which link to put to sleep first. TLS operates on a collection of traffic matrices by nature and therefore, 480 consecutive traffic matrices of week 0 of the dataset are considered. The statistical characteristics of the traffic matrices were described in detail in Section 3.5.1.

The LF and MP ETE schemes only focus on each standalone traffic matrix and therefore, 10 traffic matrices from the set of traffic matrices used to evaluate TLS were chosen to evaluate GLA on top of these schemes. The 10 traffic matrices were chosen by taking 2 traffic matrices each

from the subset of traffic matrices which has a MLU close to the Max., Min., Mean and 1<sup>st</sup> and 3<sup>rd</sup> Quartiles MLU of week 0 as specified in Table 3-3.

There are three different sets of link weights which are compared in this work: *Default*, *Interior Gateway Protocol Weight Optimizer* (IGP-WO) and *GLA*. The Default link weights are the actual link weights applied in practice.

IGP-WO contains link weights which are optimized following [52] for general load-balancing purpose only, without any energy awareness. These two link weight setting strategies are used as benchmarks to evaluate the improvement in energy-efficiency obtained by GLA based on the three common ETE schemes. GLA is run with 10 different seeds to get the average performance. As mentioned previously, GLA produces a set of Pareto-optimal solutions for each seed. Each solution candidate shows a different trade-off between energy savings and load-balancing.

As the aim is to achieve energy-efficiency without substantially sacrificing conventional traffic engineering (i.e. load-balancing) performance, it was decided to exclude all solutions which have a MLU which is  $\Psi\%$  (3% was chosen for all experiments in this chapter) above the lowest MLU given by the GLA link weights. For each seed, the best solution is then chosen by identifying the solution among the remaining solution candidates which has the lowest ratio of MLU to energy-efficiency. TLS was further evaluated by having different MLU constraints, represented by  $\alpha$  in Eq. (4.4), during off-peak time where the “reduced network topology” is applied. The different off-peak MLU constraints represent the different degrees of conservativeness by the network operator during off-peak time.

#### 4.4.2 Simulation Results

Table 4-3 and Table 4-4 demonstrate the performance of the three sets of link weights for the ETE schemes LF and MP specified in Section 4.2.2 respectively. The performance is measured in terms of the average change in Maximum Link Utilization,  $\Delta U$ , and average change in energy-efficiency,  $\Delta E$ , when compared to the results given by the Default link weights. In the case of LF, the number of sleeping links computed by GLA has increased by 16.1% while reducing the MLU by 30.7% compared to the results given by the Default link weights.

**Table 4-2: Power used due to an active link.**

$ L $	$c$ (Mbps)	$p$ (W)	$ L  \times p$ (W)
32	9953	1120	35840
2	4876	560	1120
32	2488	280	8960
8	155.2	98	784
<b><math>\Sigma</math></b>	<b>74</b>		<b>46704</b>

This shows that GLA can reduce the MLU while still achieving significantly higher energy-efficiency. IGP-WO link weight setting was not able to improve the energy-efficiency after the MLU in the network has been reduced. This is shown by the negative sign for  $\Delta E$ . It is observed that GLA performs slightly better than IGP-WO in terms of load-balancing even though GLA considers energy-efficiency at the same time. For MP, the energy-efficiency obtained by GLA has increased by 1.08% while reducing the MLU by 31% compared to the results given by the Default link weights. When IGP-WO link weights were used, it was still not able to improve the energy-efficiency when the MLU of the network is reduced.

Regarding TLS, Figure 4-3 shows that GLA can achieve a substantial improvement in energy-efficiency of 238% and 144% compared with the Default and IGP-WO link weights respectively when  $\alpha$  (in Eq. (4.4)) is set equal to the worst-case MLU given by the “full network topology”,  $U$ . Effectively,  $\alpha$  represents the worst-case MLU that can be observed during the entire off-peak operation duration of the “reduced network topology”. Similar observation is obtained when  $\alpha$  is further reduced to 65% and 60% respectively. The energy-efficiency decreases when  $\alpha$  is decreased because of the more conservative constraint for TLS. Table 4-5 shows that the high energy-efficiency obtained using GLA is not at the expense of load-balancing since the GLA values for load-balancing are lower than those for IGP-WO.

Figure 4-4 shows the actual MLU performances across the 5 days when Default, IGP-WO and GLA link weights are applied to the network. It is interesting to see that when  $\alpha$  is set equal to  $U$ , the off-peak duration of GLA is even able to cover the entire 5-day period because the difference between the peak and off-peak MLU under GLA is zero (also see Table 4-5). It is acceptable for a network to not use some links at all because this will reduce the operational costs even if the network operator has already invested capital in the network. The network operator can put the always-sleeping links back on when there is a need for extra capacity in the network. When  $\alpha$  is set below  $U$ , link sleeping can be only configured within a specific period on a daily basis. This is shown by the dark areas for the Default scenario in Figure 4-4 and for the GLA link weights in Figure 4-5 where  $\alpha$  is set equal to 60%.

**Table 4-3: Performance comparison of three sets of link weights for LF ETE scheme.**

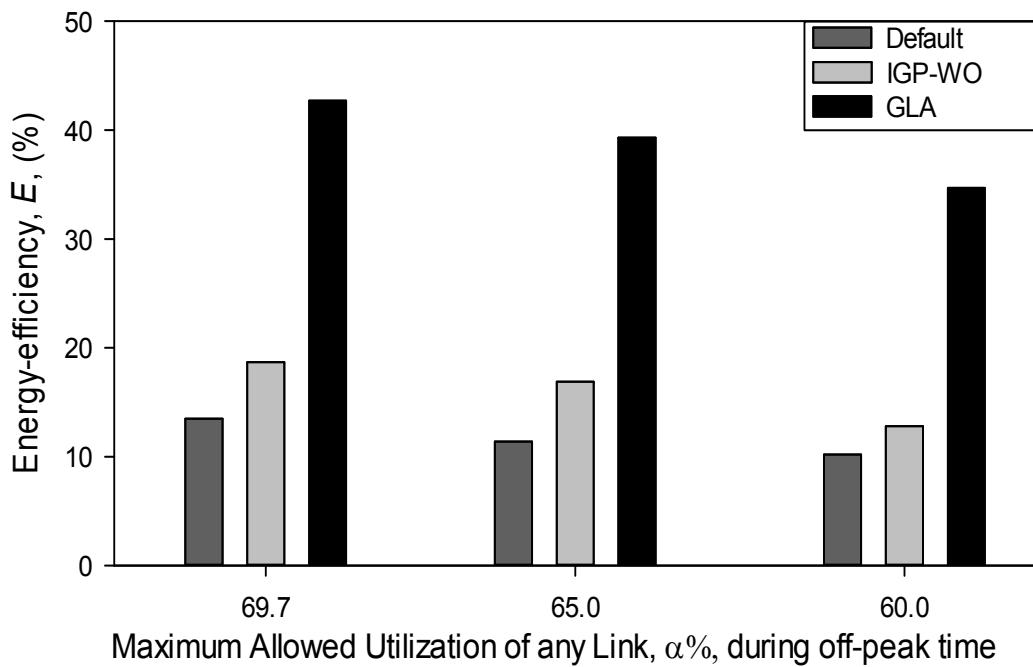
IGP-WO		GLA	
$\Delta U$ (%)	$\Delta E$ (%)	$\Delta U$ (%)	$\Delta E$ (%)
-27.1	-1.17	-30.7	16.1

**Table 4-4: Performance comparison of three sets of link weights for MP ETE scheme.**

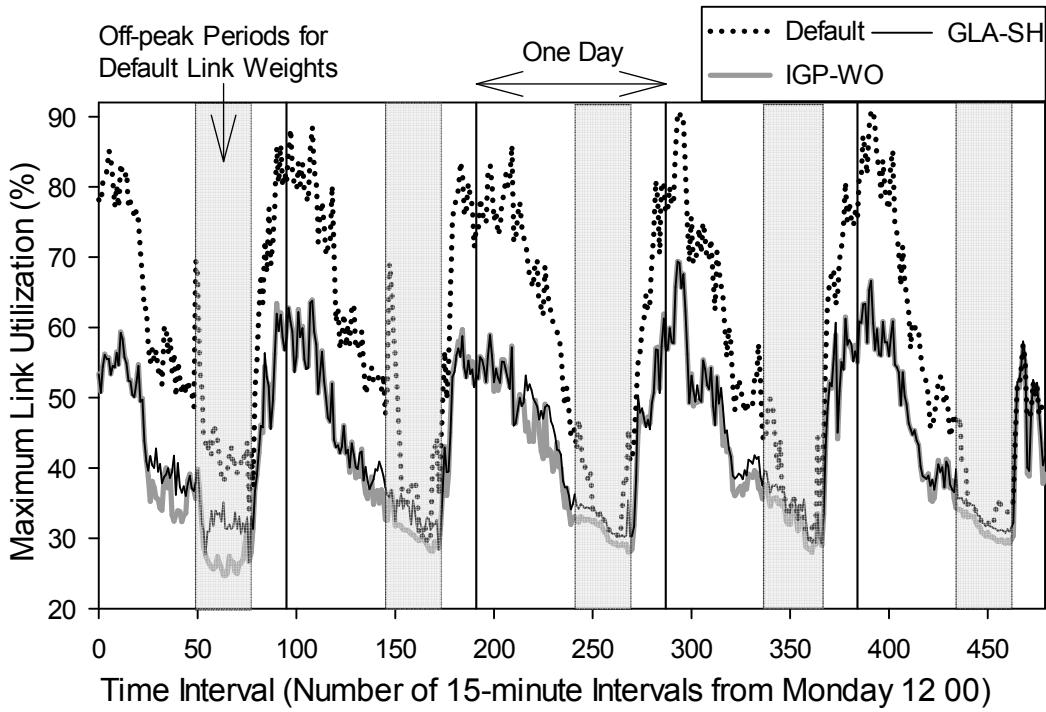
IGP-WO		GLA	
$\Delta U$ (%)	$\Delta E$ (%)	$\Delta U$ (%)	$\Delta E$ (%)
-27.1	-14.3	-31.0	1.08

**Table 4-5: Performance of GLA for the three set of link weights.**

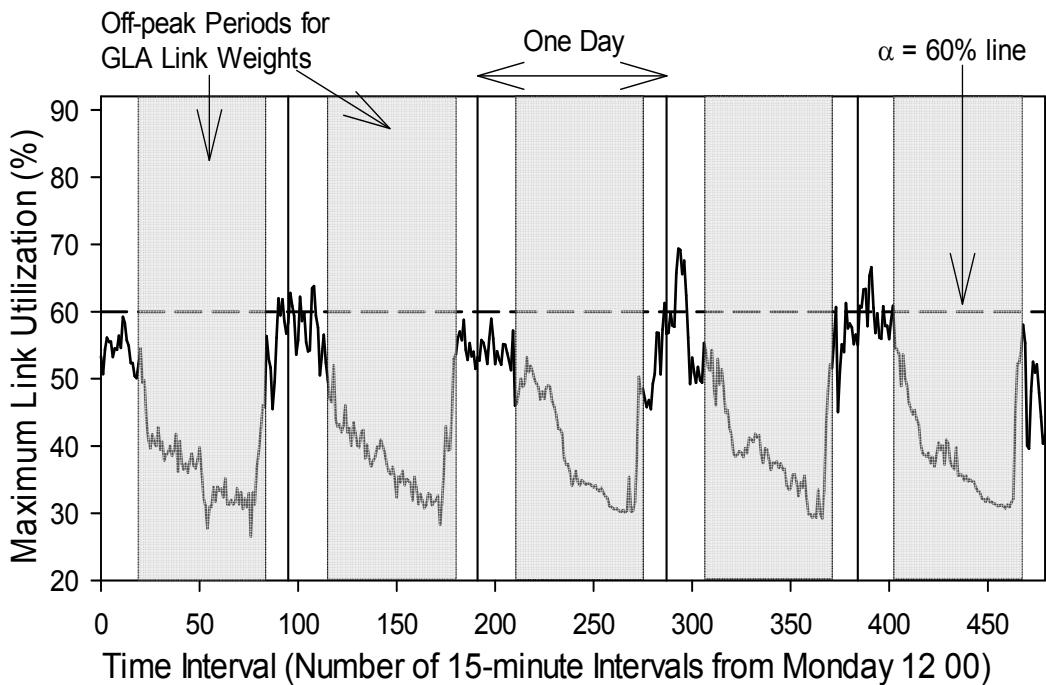
$\alpha$	Default		IGP-WO		GLA	
	$U$ (%)	$E$ (%)	$U$ (%)	$E$ (%)	$U$ (%)	$E$ (%)
69.7	90.9	13.5	70.1	18.7	69.7	42.7
65.0	90.9	11.1	70.1	16.9	69.5	39.3
60.0	90.9	10.2	70.1	12.8	69.6	34.7



**Figure 4-3: Energy-efficiency of TLS using the different sets of link weights.**



**Figure 4-4:** The MLU variation across 5 days for Default, IGP-WO and GLA link weights when  $\alpha$  is set equal to  $U$ .



**Figure 4-5:** The MLU variation across 5 days for GLA link weights when  $\alpha$  is set equal to 60, which is below  $U$ .

## 4.5 Enhanced GLA for TLS

### 4.5.1 Solution-enhancement Heuristic

The performance of GLA for TLS can be improved further if GLA is further customized for TLS through the use a Solution-enhancement Heuristic (SH). This enhanced version of GLA is called *Green Load-balancing Algorithm with Solution-enhancement Heuristic* (GLA-SH) and is designed to run at the end of each iteration of GLA. SH operates on the best solution in the population at the end of each iteration of GLA. The best solution is determined by a single aggregated objective function represented by the ratio of Maximum Link Utilization to energy-efficiency as given by Eq. (4.1) and (4.2) in Section 4.2.1 respectively. Figure 4-6 shows the flowchart of the operation of SH. The chromosome (i.e. solution candidate) selected for improvement is first evaluated to find the size of the off-peak duration. In TLS, the energy-efficiency is effectively represented by the number of sleeping links multiplied by the off-peak duration.

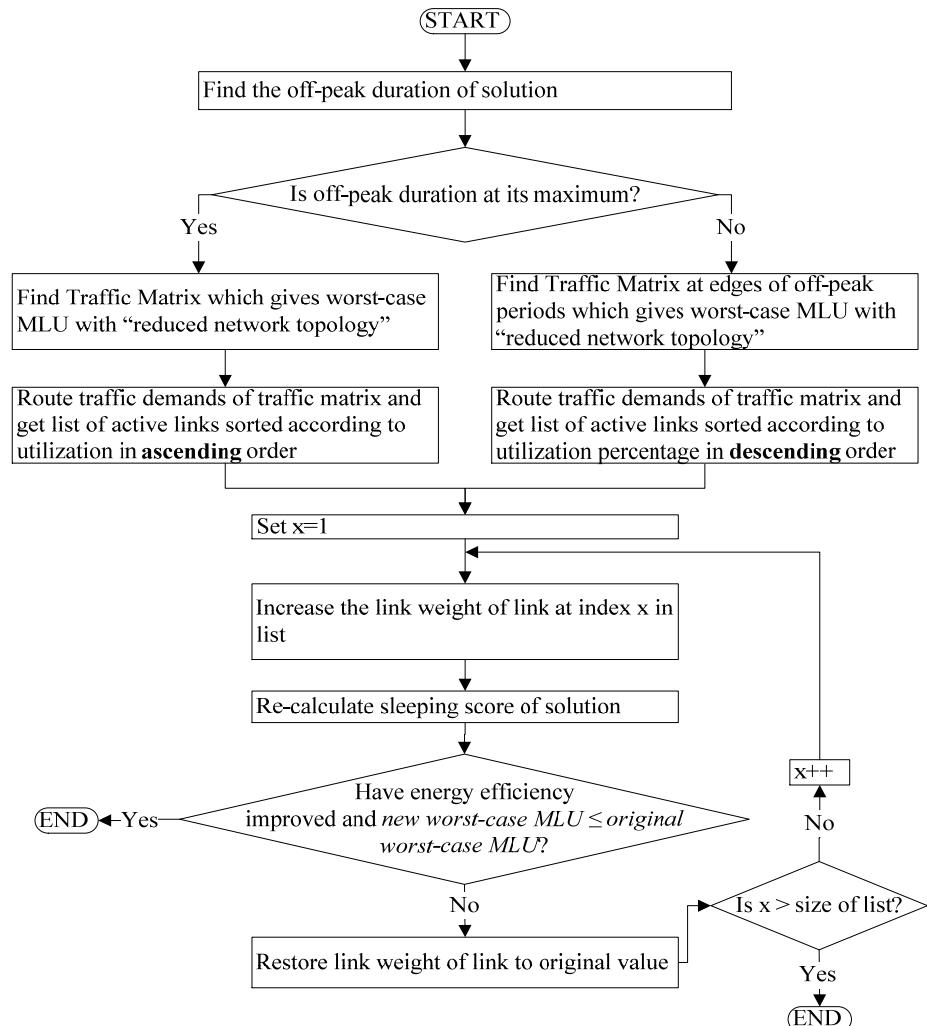


Figure 4-6: Flow chart describing the operation of the Solution-enhancement Heuristic.

If the length of the off-peak duration is at its maximum, the duration cannot be further increased to improve the energy-efficiency. The only option is to increase the number of sleeping links instead. This involves diverting load from certain active links in order to enable them to sleep. This diversion of load is done through the increase of the IGP link weights of certain links. SH first identifies the traffic matrix which gives the worst-case MLU when its traffic demands are routed with the off-peak “reduced network topology”. The next step involves the creation of a list of active links ranked in ascending order according to the utilization of the links. The first link in the list is the least utilized in the network and therefore, it is easier to shift its load to other links. This link will be first chosen to have its link weight increased. In this work, the link weights were increased by 20% but this value can be changed depending on network scenario. After each link weight increase, the modified chromosome is re-evaluated to see if the energy-efficiency has improved and the worst-case MLU has remained the same or has been reduced. If these criteria are met, SH stops and returns the improved chromosome to the population where it will replace the currently worst chromosome. In the case of the criteria not being met, the link weight of last modified link is restored to its original value and the next link in the list undergoes link weight increase. This process continues until either all the links in the list have been tested or improvement of the chromosome has been successful.

If the size of the off-peak duration is not at its maximum, the energy-efficiency can be improved through the increase of the length of the duration of the off-peak period. The traffic matrices at the edges of the off-peak periods are first evaluated according to the off-peak “reduced network topology” to see which one gives the worst-case MLU. The identified traffic matrix is the one which has most likely caused the off-peak duration to be small because the maximum allowable utilization of any link in the network has been reached. The traffic demands of the identified traffic matrix are then routed and a list of active links is created. This time, the list is sorted in descending order according to utilization percentage. This is because the MLU can be reduced to allow the off-peak duration to increase for more energy-efficiency gains. This is achieved by shifting load from the most utilized link by increasing its link weight. The same iterative process of link weight increase is then performed in the same manner as in the previous case. Chromosomes which cannot be improved by SH are tracked so that SH does not run on them again and the next best chromosome is used as candidate for improvement by SH.

It is worth noting that it is possible to introduce solution enhancement heuristics for other ETE algorithms as well. These solution enhancement heuristics need to be specifically designed by taking into account their own working mechanisms in question. The development of such heuristics is however outside the scope of this thesis.

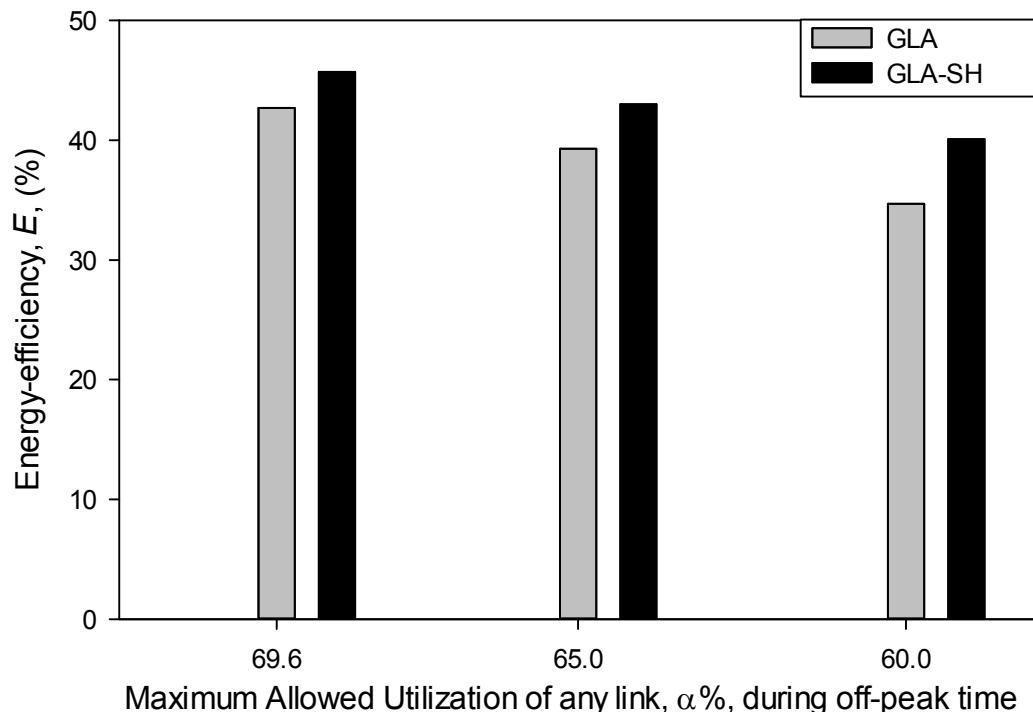
#### 4.5.2 Evaluation

The GLA-SH was evaluated using the same network scenario as described in Section 4.4.1. The results from Table 4-6 show that GLA-SH can increase the energy-efficiency of TLS by 7% compared to plain GLA when  $\alpha$  is set equal to  $U$ .

The improvement when  $\alpha$  is set to 65% and 60% is 9.41% and 15.6% respectively. These results show that the Solution-enhancement heuristic can improve the energy-efficiency of TLS by customizing the operation of GLA. The reason why GLA-SH is able to improve on plain GLA is because the SH allows GLA-SH to escape from a local optimal solution. Genetic algorithms, on which GLA is based on, are known to be trapped in local minima and therefore, are not guaranteed to converge to the true global optimal [71].

### 4.6 Maximum Packet Delay (MPD) Performance of the existing ETE Schemes

The end-to-end Maximum Packet Delay (MPD) of operational networks has become increasingly important due the popularity of real-time applications such as Voice-over-IP (VoIP) and video streaming [39].



**Figure 4-7: Comparison of energy-efficiency between GLA and GLA-SH.**

**Table 4-6: Performance comparison between GLA and GLA-SH.**

$\alpha$	GLA		GLA-SH	
	$U$ (%)	$E$ (%)	$U$ (%)	$E$ (%)
69.6	69.7	42.7	69.6	45.7
65.0	69.5	39.3	69.5	43.0
60.0	69.6	34.7	69.5	40.1

These applications normally require a bounded MPD in order to provide end-users with good Quality-of-Experience. The three ETE schemes, described in Section 4.2.2: LF, MP and TLS, are oblivious to packet delay and therefore, the packet delay may substantially increase when links are put to sleep for energy savings. In this section, a comprehensive analysis of the MPD performance of the three different ETE schemes on the three different sets of link weights: Default, IGP-WO and GLA is presented.

Table 4-7 shows the maximum, average and minimum increase in MPD when the LF ETE scheme is applied for the Default, IGP-WO and GLA link weights. The maximum packet delay was measured in the same way as described in Section 3.5.2. The general observation is that Default link weights perform the best in terms of MPD performance while the GLA link weights scenario has the worst result. There are two main reasons behind this observation: first, the Default link weights have more freedom in their routing because they do not optimize load-balancing as GLA does. The second reason is that GLA puts more links to sleep compared to the Default link weights as shown in Table 4-3 in Section 4.4.2 where the LF ETE scheme measures directly the number of links put to sleep (see Eq. (4.5)). This leads to the network topology being less connected, forcing packets to take longer paths to reach their final destinations. GLA has also worse MPD performance compared to IGP-WO because even though the load-balancing performance is similar, more links are put to sleep by GLA compared to IGP-WO. The IGP-WO and GLA columns in Table 4-7 indicate that there can be large variations in MPD performance given different traffic matrices. This is due to the fact that the LF ETE scheme is putting links to sleep with the only constraints of maintaining full network connectivity and not violating the MLU threshold  $\alpha$ . Hence, the LF ETE scheme does not consider any packet delay requirement. Table 4-8 further shows the same general trend in MPD performance for the MP ETE scheme as for the LF ETE scheme based on the three sets of link weights: Default, IGP-WO and GLA. The reasons for this observation are the same as for the LF ETE scheme.

As mentioned in Section 4.2.2.3, TLS uses a full and a reduced network topology for peak and off-peak time operation respectively.

Table 4-8 shows the MPD performance of TLS when the Default link weights are used. Upon analysis of the reduced network topologies which were generated with the Default link weights and various values of  $\alpha$ , it was found that all of them are the same and have 33 sleeping links each. Since the reduced network topology for the various values of  $\alpha$  is the same (but have different off-peak durations), the value of the maximum increase in MPD is the same.

Another observation is that the maximum MPD values for the reduced network topologies are smaller compared to the values in Table 4-7. Further analysis of the number of sleeping links in each ETE scheme shows that on average, 39.6 links are sleeping in the LF ETE scheme compared to 33 links for TLS ETE scheme. This larger number of sleeping links reduces the connectivity of the network, which explains the increase of the MPD. TLS has a lower number of sleeping links because it trades off sleeping links for a longer off-peak duration to maximize energy-efficiency with only two network topologies.

Table 4-9 also shows the maximum increase in MPD when IGP-WO link weights are used for both the full and reduced TLS network topologies. The main observation is that the MPD is very large. As mentioned previously, the IGP-WO algorithm is oblivious to packet delay and optimizes the load-balancing performance only.

**Table 4-7: Increase in maximum packet delay for the reduced network topology due to the Default, IGP-WO and GLA link weights for LF ETE scheme.**

Increase in Maximum Packet Delay	Value (%)		
	Default	IGP-WO	GLA
Maximum	2.56	80.0	81.5
Average	0.627	19.8	44.2
Minimum	0.149	0.635	2.58

**Table 4-8: Increase in maximum packet delay for the reduced network topology due to the Default, IGP-WO and GLA link weights for MP ETE scheme.**

Increase in Maximum Packet Delay	Value (%)		
	Default	IGP-WO	GLA
Maximum	25.39	88.0	83.6
Average	6.99	55.3	78.9
Minimum	2.97	0.713	55.8

**Table 4-9: Increase in maximum packet delay due to the Default, IGP-WO and GLA-SH link weights for TLS ETE scheme.**

$\alpha$	Maximum increase in Maximum Packet Delay (%)					
	Default		IGP-WO		GLA-SH	
	Full NT	Reduced NT	Full NT	Reduced NT	Full NT	Reduced NT
69.6	0	0.336	54.1	54.1	77.3	77.3
65.0	0	0.336	54.1	54.1	77.0	80.0
60.0	0	0.336	54.1	54.1	77.0	80.0

Table 4-8 shows GLA-SH link weights have the worst MPD performance among the three sets of link weights. This is because GLA-SH puts a significantly larger number of links to sleep and achieves a better load-balancing performance compared to the other scenarios as shown in Table 4-3 and Table 4-5.

As mentioned in Section 4.2.2.3, TLS uses only a single reduced network topology all the time when  $\alpha$  is set to 69.7 for the GLA link weights and therefore, the MPD performance remains the same during both peak and off-peak time. For values of  $\alpha$  of 65 and 60, two network topologies are used for peak and off-peak operation. The reduced network topology used during off-peak time has some links put to sleep compared to the full network topology which has no sleeping links. As a result, there is a higher probability for the MPD to increase when the reduced topology is used.

## 4.7 Improvement to current ETE Schemes for Maximum Packet Delay Constraint

In the previous section, it was indicated that the MPD performance of the three plain ETE schemes can vary substantially with the three sets of link weights. The current ETE schemes do not have a deterministic way of ensuring that a bounded MPD performance is maintained. Therefore, there is no guarantee that the MPD performance will remain within the tolerance limits of network operators. In light of this observation, it is desired to further extend the current ETE schemes so as to guarantee a minimum MPD performance while maintaining the load-balancing and energy-efficiency performance of the schemes.

To take into account this new requirement, the MPD constraint in Eq. (4.9) is added to the problem formulation in Section 4.2.1 where  $D_{new}$  is the MPD of the new network topology after it has been modified for load-balancing and/or energy-efficiency,  $D_{original}$  is the MPD of the original network topology and  $\Omega$  is the maximum allowed increase ratio in MPD as defined by the

network operator. The new MPD-aware ETE schemes have the same algorithmic complexity as their plain counterparts since the MPD of a reduced network topology can be calculated during the calculation of the forwarding tables.

$$D_{new} \leq \Omega \times D_{original} \quad \text{with } \Omega \geq 1.0 \quad (4.9)$$

In order to enforce the new MPD constraint in Eq. (4.9), the way through which the three ETE schemes put links to sleep needs to be modified. The modification is that in addition to checking if the MLU and full connectivity constraints are maintained when each candidate link is considered to be put to sleep, the MPD performance of the resulting topology also needs to be examined. If the resulting MPD is higher than the upper bound, the link under consideration should not be put to sleep. The plain ETE schemes, LF, MP and TLS, which are modified to include the MPD constraint, are named LF-D, MP-D and TLS-D respectively.

## 4.8 Performance Evaluation of ETE Schemes with Maximum Packet Delay Constraint

The three new MPD-aware ETE schemes, namely LF-D, MP-D and TLS-D, are evaluated by using the same network scenario as described in Section 4.4.1 and compare the load-balancing and energy-efficiency performance when the Default and GLA link weights are used. The IGP-WO link weights were not used during the evaluation because they give an MPD value which is very high. Specifically, the newly introduced extra delay due to the optimization of the link weights for load-balancing only is substantially higher even if no links are put to sleep as shown in Table 4-10.

### 4.8.1 Performance Comparison between the Original and MPD-aware ETE Schemes

In this section, the change in load-balancing and energy-efficiency performance is compared between the original and the MPD-aware ETE schemes. The chosen value of  $\Omega$  is 1 for all the remaining experiments in this chapter. Table 4-11 shows that the energy-efficiency performance decreases when the LF-D and MP-D ETE schemes are applied to the Default link weights when compared to the original schemes. This decrease in performance is due to a smaller number of links being put to sleep, so that the network topology remains sufficiently connected to keep the MPD within the set constraint. Table 4-12 shows a decrease in both the energy-efficiency and load-balancing performance when GLA is used with the LF-D scheme. These reductions are due to the new constraint on the MPD which reduces the diversity of paths available in the network after link sleeping. For the MP-D scheme, there is actually a slight improvement in the load-

balancing but this comes at the expense of a greater decrease in the energy savings. Table 4-13 shows a slight increase in the load-balancing performance of the GLA-SH link weights for the TLS-D ETE scheme but the energy-efficiency performance decreases.

#### **4.8.2 Performance Comparison between Default and GLA link weights for MPD-aware ETE Schemes**

In this section, a direct comparison study, in terms of load-balancing and energy-efficiency, between the Default and the GLA link weights for the MPD-aware ETE schemes is presented. Therefore, it is possible to get a moderate gain in energy-efficiency if solutions which do not have close-to-optimal load-balancing performance are chosen.

Table 4-15 shows that it is possible to use GLA-SH to achieve significant improvement in both load-balancing and energy-efficiency performance over the Default link weights for the TLS-D scheme.

**Table 4-10: Increase in MPD for the full network topology due to the IGP-WO link weights compared to the Default ones.**

Increase in MPD	Value (%)
Maximum	54.1
Average	53.8
Minimum	53.0

**Table 4-11: Change in energy-efficiency when LF-D and MP-D, rather than LF and MP, are applied to the Default link weights.**

ETE Scheme	$\Delta E(\%)$
LF-D	-3.52
MP-D	-2.52

**Table 4-12: Change in the improvement of the load-balancing and energy-efficiency when LF-D and MP-D, rather than LF and MP, are applied to the GLA link weights.**

ETE Scheme	$\Delta\Delta U(\%)$	$\Delta\Delta E(\%)$
LF-D	-0.312	-6.77
MP-D	0.776	-7.53

**Table 4-13: Performance improvement of the load-balancing and energy-efficiency when TLS-D, rather than TLS, is applied to the GLA-SH link weights.**

$\alpha$	GLA-SH	
	$\Delta U$ (%)	$\Delta E$ (%)
69.4	0.190	-5.48
65.0	0.259	-5.50
60.0	0.236	-15.6

Table 4-14 shows the change in load-balancing and energy-efficiency performance when the GLA link weights are used, instead of the Default ones, for the LF-D and MP-D ETE schemes. For the LF-D scheme, GLA is able to improve both the load-balancing and the energy-efficiency performance over the Default link weights. For the MP-D scheme, GLA is only able to improve the load-balancing performance at the expense of the energy-efficiency. When GLA converges, a final population of solutions, which contains a number of Pareto-optimal solutions, is obtained. Pareto-optimal solutions are solutions, which arise in multi-objective optimization, where one objective cannot be improved at the expense other objectives. In short, Pareto-optimal solutions offer different trade-offs in terms of load-balancing and energy savings. If the energy savings performance is prioritized over the load-balancing performance for the MP-D scheme, it is possible to get values of -4.11% and 5.29% for  $\Delta U$  and  $\Delta E$  respectively.

**Table 4-14: Performance improvement when GLA link weights are used for the LF-D and MP-D ETE scheme compared to Default link weights.**

ETE Scheme	$\Delta U$ (%)	$\Delta E$ (%)
LF-D	-30.9	12.2
MP-D	-30.5	-4.11

**Table 4-15: Performance improvement when GLA-SH link weights are used for the TLS-D ETE scheme compared to Default link weights.**

$\alpha$	GLA-SH	
	$\Delta U$ (%)	$\Delta E$ (%)
69.4	-23.5	46.8
65.0	-23.6	185
60.0	-23.6	167

## 4.9 Summary

In this chapter, a *Green Load-balancing Algorithm* (GLA) is proposed which intelligently optimizes IGP link weights in IP backbone networks in order to improve both the load-balancing and energy-efficiency of existing ETE schemes. The most important finding from this chapter is that it is possible to optimize the link weights so that significant energy-savings can be obtained while preserving near-optimal load-balancing. This is an important result since network operators do not have to make a hard choice between energy-efficiency and load-balancing. Therefore, the previous assumption that significant energy-savings cannot be done when load-balancing is done is not correct. Moreover, the design of GLA enables the network operator to do a tradeoff between load-balancing and energy-efficiency based on the policy that it has.

GLA uses a customized multi-objective genetic algorithm to identify the optimized solutions. In addition, two new custom mutation and crossover operators have been designed to improve the performance of the genetic algorithm by enabling the solution space to be searched more efficiently. In order to be able to draw generic conclusions about the energy-efficiency and load-balancing performance of GLA, three diverse existing ETE schemes, LF, MP and TLS, were selected to extensively evaluate the ability of GLA to jointly optimize the load-balancing and energy-efficiency performance of these schemes. Simulations based on the European academic network GÉANT and its real traffic matrices were used to evaluate GLA. GLA was shown to improve the energy-efficiency of LF, MP and TLS by 16.1%, 1.08% and 216% respectively compared to the Default link weight setting scenario. This improvement has been achieved while maintaining near-optimal load-balancing performance as shown through a comparison with IGP-WO. An enhanced version of GLA, GLA-SH, was also designed specifically for TLS. GLA-SH was able to improve the energy-efficiency of TLS by 239% compared to the Default link weight setting, while maintaining near-optimal load-balancing performance.

In the second part of this chapter, the Maximum Packet Delay (MPD) performance of the same three ETE schemes was analyzed. It was found that the ETE schemes often exhibit excessive MPD because they are completely oblivious to the increase in packet delay when they put links to sleep to save energy and/or when they optimize the link weights to improve the load-balancing in the network. In light of this observation, the plain ETE schemes were modified to take into account the MPD performance and the load-balancing and energy-efficiency performance of the new MPD-aware ETE were evaluated. It was found that it was still possible to save significant energy while achieving near-to-optimal load-balancing in most scenarios and still respecting a bound on the maximum packet delay.

The performance of the reduced network topology, obtained by GLA on top of TLS, during single link failures can be guaranteed by using an extended version of TLS called *Time-driven Link*

*Sleeping with Single Link Failure Protection* (TLS-SLFP) which was described in detail in Chapter 3. GLA can operate with no modification on top of TLS-SLFP and this will be studied in future work along with extensions to the LF and MP schemes to support single link failures.

In summary, GLA can be regarded as a very promising approach which is able to further enhance the performance of ETE algorithms while maintaining at the same time the capability of the produced network configurations towards supporting traditional traffic engineering objectives such as load-balancing and MPD.

## Chapter 5

# 5 Leveraging MPLS Backup Paths for Distributed Energy-aware Traffic Engineering<sup>4</sup>

## 5.1 Introduction

In sharp contrast with the two previously-designed ETE schemes TLS and GLA; this chapter introduces a new online and fully-distributed ETE scheme called *Green Backup Paths* (GBP). Online ETE schemes are important for networks which have an unpredictable traffic behavior, such as the one shown in Figure 1-1. This type of network needs an ETE scheme which can quickly capture the unusual traffic patterns and optimize the network accordingly so as to preserve the energy-efficiency of the network as well as the other network objectives such as load-balancing when the traffic demands vary in the network. Most online ETE schemes are made up of two main components. The first component is a traffic monitoring component which can accurately and precisely measure the traffic in the whole network. In addition, this component needs to be able to inform all relevant decision-making entities about the changes in the traffic demands in the network. The second component of an online ETE scheme will then use the traffic information from the first component to make optimized network routing decisions which will improve the overall energy-efficiency of the network. Since traffic demands can rapidly change in some networks, it is important for online ETE schemes to converge quickly. One way of achieving this is to enable the different ETE decision-making entities in the network to concurrently make interference-free decisions.

---

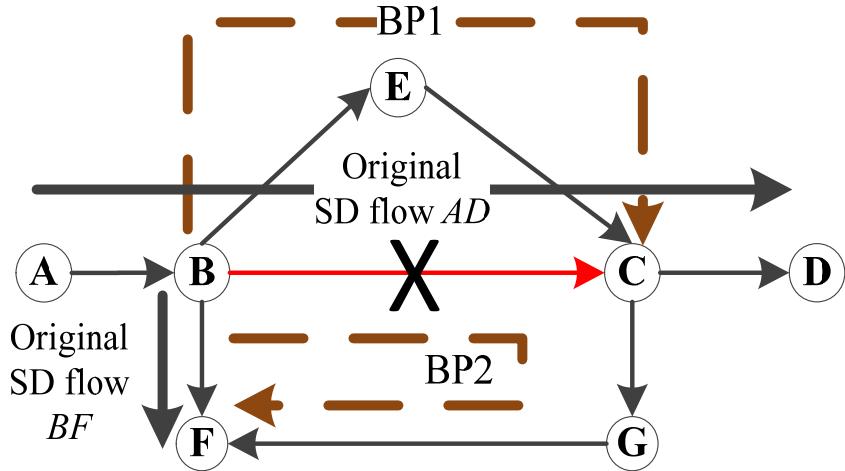
<sup>4</sup> The content of this chapter is based on the following publications:

- 1) F. Francois, N. Wang, K. Moessner, and S. Georgoulas, "Leveraging MPLS Fast ReRoute Paths for Distributed Green Traffic Engineering", in Proc. of 2013 IEEE International Workshop on Quality of Service (IWQoS), June 2013 (**Short Paper and Poster Session**).
- 2) Frederic Francois, Ning Wang, Klaus Moessner, Stylianos Georgoulas and Ricardo Schmidt, "Leveraging MPLS Backup Paths for Distributed Energy-aware Traffic Engineering," IEEE Transactions on Network and Service Management (TNSM) (**under review**).

Nowadays, it is common for backbone operators to use Multi-Protocol Label Switching (MPLS) to explicitly route traffic between the different Source-Destination (SD) pairs in their networks. These backbone networks are protected against single link failures through the use of pre-installed backup paths [72]. A backup path is used to divert the affected traffic away from the protected link when it fails. The route taken by the backup path is usually the shortest path between the head and tail router of the protected link but without traversing the protected link. In such local protection, the failure recovery is handled only by the head router of the protected link, and none of the other remote routers needs to be aware of the failure if it occurs. The illustrative network topology in Figure 5-1 is used to demonstrate how the traffic on a link is protected against the failure of the link by a pre-installed backup path. For example, if the link  $B \rightarrow C$  fails, the head router  $B$  of the link will divert the flow from  $A$  to  $B$  onto the pre-installed backup path  $BP1$  ( $B \rightarrow E \rightarrow C$ ) to avoid packet loss.

GBP improves the power-efficiency of networks by diverting traffic away from protected links onto the backup paths in an intelligent manner, so that the protected links can go to sleep. The links of the backup paths are already active in the network since they carry their own traffic and therefore, the diversion of traffic from protected links onto them do not increase the power consumed by these links since the amount of traffic on a link does not currently affect much its power consumption [8] as explained in more detail in Section 2.2. GBP directly uses the existing backup paths for power savings while not impairing the ability of these paths to protect against single link failures. For each link being considered for sleeping, this is achieved by carefully diverting traffic onto its backup path while making sure that any single link failure during the sleeping period will not cause traffic overloading at any link on that backup path. In addition, GBP also considers the traditional traffic engineering function of resilience of the network against potential traffic upsurges by not causing any link to become overloaded during any of its operations. On the contrary, GBP actively attempts to reduce the traffic on overloaded links so that the peak link utilization decreases in the network. Hence, a second objective of GBP is to increase the resilience of the network against potential traffic upsurges, which is in addition to its other objective of power savings.

The key novelty of GBP is the exploitation of existing failure-protection backup paths for the dual purpose of power savings and protection against link failures. This brings the main benefit of achieving power-efficiency without installing any other paths in addition to the backup paths, which are needed anyway for failure protection. Indeed, GBP differs significantly from most other existing online traffic engineering schemes (e.g. [11][32][38]) which target either power savings or link failure protection but not both. Furthermore, GBP considers Quality-of-Service (QoS) by *actively* avoiding the use of excessively long backup paths for power savings so as to avoid substantial packet delays, but allows the use of such paths for handling link failures.



**Figure 5-1: Basic network topology to illustrate how links are protected in MPLS backbone networks.**

An additional advantage of GBP is its fast path manipulation. As will be shown later, multiple routers can make concurrent conflict-free decisions at the same time thanks to their knowledge of interference relationships with each other. Moreover, GBP uses only a single path to route each SD flow and therefore, avoids packet reordering linked with multi-path routing.

In order to evaluate the performance of GBP, the publicly-available topology and real traffic matrices of two academic backbone networks were used, namely GÉANT and Abilene. It was observed that GBP can achieve significant power savings which are always within 15% of the theoretical upper bound. This result was achieved without any increase in the peak MLU of the network as a trade-off. In addition, the ability of GBP to reduce the MLU in the network was also evaluated. According to the evaluation results, GBP was able to even significantly reduce the MLU in the case of GÉANT where there is a large diversity of paths and enough spare capacity. Furthermore, single link failures were simulated in the network and it was observed that the use of GBP did not increase the post-failure peak MLU. In addition, the increase in the maximum packet delay due to the use of the backup paths was also found to be minimal and acceptable. Therefore, QoS constraints linked with delay can be met while performing online sleeping reconfigurations through the necessary traffic diversion by GBP.

The rest of this chapter is organized as follows: in Section 5.2, the problem formulation for the GBP scheme is described in detail. In Section 5.3, an overview of GBP is first presented and then an extensive description of each component of GBP is provided along with flow charts to show each process of every component of GBP. In Section 5.4, the results from the evaluation of GBP on the GÉANT and Abilene topologies are presented. Finally in Section 5.5, this chapter is concluded with the key findings.

## 5.2 Problem Formulation

**Table 5-1: Definition of symbols for GBP.**

Variable	Description
$\alpha$	Maximum allowable utilization of logical link
$b_{ij}^{sd}$	Specifies if a logical link from router $i$ to $j$ is used to route traffic from router $s$ to $d$ . A value of “1” means the logical link is used, otherwise it is “0”.
$c_l$	Capacity of logical link $l$
$f_{ij}^{sd}$	Traffic demand from router $s$ to $d$ that traverses logical link from router $i$ to $j$
$f_l$	Total traffic demand on logical link $l$
$G(R, L)$	Directed graph with $R$ being set of routers and $L$ being set of logical links
$p_l$	Power consumed by an active physical link in logical link $l$
$t^{sd}$	Traffic demand from router $s$ to $d$
$y_l^o$	Number of physical links in logical link $l$ which are in sleep mode

Nowadays, a logical link between router pairs in networks is usually made up of a bundle of physical links [20][27]. Such a strategy reduces the complexity in upgrading network capacities by adding new physical links to the existing bundle. If traffic demands are lower than the capacity of the whole bundle, power savings can be achieved by putting *unused* physical links to sleep. In addition, the line card connected to a physical link can have the opportunity to sleep when the physical link is put to sleep. Sleeping line cards are the major source of power savings in green networks because they contribute up to 42% of the total power consumption of a backbone router [11]. Putting part of a logical link to sleep can be viewed as a form of rate adaptation, which is similar to what has been developed for green Ethernet [16].

The actual online optimization within each periodical GBP operation cycle can be expressed as:

$$\begin{aligned} & \text{minimize } f_l - \frac{\alpha}{100} c_l \quad \forall l \text{ where } f_l > \frac{\alpha}{100} c_l \\ & \text{with } \alpha = [0, 100] \end{aligned} \quad (5.1)$$

$$\text{maximize } \sum_{l=1}^{|L|} (y_l^o \times p_l) \quad (5.2)$$

subject to:

$$\sum_{j=1}^{|R|} b_{ij}^{sd} - \sum_{j=1}^{|R|} b_{ji}^{sd} = \begin{cases} 1 & \forall s, d, i = s \\ -1 & \forall s, d, i = d \\ 0 & \forall s, d, i \neq s, d \end{cases} \quad (5.3)$$

$$\sum_{j=1}^{|R|} b_{ij}^{sd} f_{ij}^{sd} - \sum_{j=1}^{|R|} b_{ji}^{sd} f_{ji}^{sd} = \begin{cases} t^{sd} & \forall s, d, i = s \\ -t^{sd} & \forall s, d, i = d \\ 0 & \forall s, d, i \neq s, d \end{cases} \quad (5.4)$$

$$f_l < \frac{\alpha}{100} c_l \quad \forall l \text{ with } \alpha = [0, 100] \quad (5.5)$$

Equation (5.1) is the first objective of GBP which is to minimize the Maximum Logical Link Utilization (MLLU) in the network so that the network is more resilient to traffic upsurges because of the more balanced load. The MLLU was referred to as the MLU before the introduction of the concept of bundle links in this section. Equation (5.2) represents the second objective of GBP which is to maximize the total amount of power saved in network operations. This is represented by the sum (over all logical links of the network) of the product of the number of sleeping physical links in a logical link and the power consumed by the physical links if they were left active. Equation (5.3) is the constraint which enforces a single path to be taken by all traffic which has the same source and destination. Equation (5.4) is the conventional flow conservation constraint. Equation (5.5) prevents a logical link from being loaded above the threshold,  $\alpha$ , due to the operation of GBP. Moreover, GBP does not use backup paths for power saving if their path length (delay) is too long but they will still be used for link failure protection.

The problem of maximizing the number of physical links which can go to sleep while respecting the above constraints has already been proven to be NP-hard in [11]. Therefore, a computationally-efficient heuristic, called *Green Backup Paths* (GBP), which can be applied in a network in an online and distributed fashion without requiring significant modifications to existing network protocols is presented.

## 5.3 Green Backup Paths (GBP)

In this section, an overview of GBP is first presented followed by an in-depth description of all its different components. Table 5-2 acts as a point-of-reference for the name and description of all the notions used during the description of the operation of GBP.

### 5.3.1 Scheme Overview

The proposed GBP scheme consists of two distinct operational components, namely the *offline* and the *online* components. The offline component identifies the eligible backup paths for GBP operations. This is done based on the delay (length) characteristics of the paths. Network operators can obtain the delay of a path by using the physical length of the path and the speed of the path medium.

**Table 5-2: Description of all notions used in GBP.**

Name	Description
<i>b_flows_list</i>	List of all SD flows that were diverted onto the logical link by GBP and the logical link has the same head router as the protected link
<i>conflict_links_set</i>	List of all logical links which are already in use by Token Holding (TH) routers in the current Multiple TH Links Selection iteration
<i>IRLL</i>	Each logical link has an Interference-Risk Links List (IRLL) to store the logical links whose spare capacity must not be modified when the logical link is undergoing offloading by GBP
<i>normal_links_list</i>	All logical links not in <i>priority_links_list</i>
<i>priority_links_list</i>	List which contains all the logical links that have their utilization above a pre-defined threshold $\alpha$
<i>p_flows_list</i>	List of all SD flows that normally use the logical link
<i>s_flows_list</i>	List of all SD flows on the logical link which are not in the <i>p_flows_list</i> and <i>b_flows_list</i> .
TH link	A logical link which has been selected by GBP for part of its traffic to be offloaded to an alternate route
TH router	A router which is the head router of a TH link and therefore, it is the network device which is responsible to attempt the traffic reroute away from the TH link

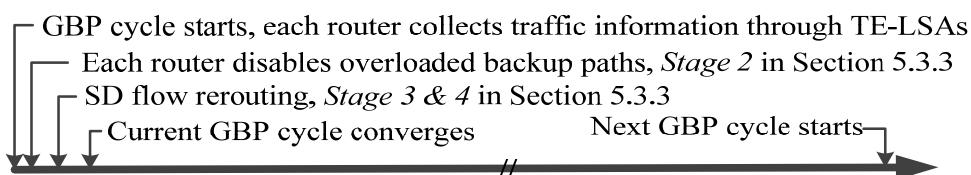
The offline component also identifies an *Interference-Risk Links List (IRLL)* for each logical link in the network. The IRLL for a logical link contains all the logical links that can be potentially affected by this link if it is offloaded. The IRLLs are essential for GBP to be able to concurrently and independently offload multiple logical links without any conflict. The details of how the IRLLs are obtained and used will be described in Section 5.3.2.2 and 5.3.3.3. If the alternative path for a logical link introduces substantially longer delay, the network operator may choose not to offload such a logical link for power savings and MLLU reduction purposes. Such logical links are also identified in the offline phase. The IRLL and eligibility for GBP operations for each logical link are distributed to the routers only once since this information is static as long as the network topology is not changed.

The second component of GBP performs an online optimization by using a heuristic to periodically divert traffic away from logical links by activating/deactivating backup paths in order to optimize both the power savings and the MLLU within the network. The periodicity of the online operation can be determined by the network operator as a trade-off between the overhead of monitoring the network and the need to detect any significant changes in the network traffic condition.

Figure 5-2 shows an illustration of the online GBP operation cycle. When a new GBP cycle starts, each router collects information about the traffic conditions by receiving the Traffic Engineering-Link State Advertisements (TE-LSAs) [73] that are broadcast by every router in the network. Based on this traffic information, each router can then calculate whether any of its directly attached logical links can successfully offload part of their traffic onto alternative paths to save power and/or reduce the MLLU. If sufficient traffic offloading is achieved, one or more physical links in the concerned logical link can go to the sleep mode.

At the beginning of a GBP optimization cycle, each router needs to check if there are any logical links which have become overloaded, i.e. do not comply with the constraint in Eq. (5.5) because of the sleeping reconfigurations in previous GBP optimization cycles. This may happen due to the increased volume of incoming traffic since the last traffic monitoring observation. In case such logical links are identified, the associated router will wake up some sleeping physical links in the relevant protected logical links and restore the currently diverted traffic back to the protected logical link(s). As a result, the previously active backup path is deactivated and traffic is no longer diverted on its links. After each router deactivates these overloaded backup paths, they wait for a settling period and then broadcast a new TE-LSA to notify all other routers about the new state of their logical links.

Each router then continues the online decision process by collecting the new TE-LSAs and updating the list of logical links. The list of logical links is always sorted at each router such that all routers have an identically ordered list. Each router goes through the list and selects the Token Holding (TH) links that can be concurrently offloaded without interfering with each other. A TH link is a logical link selected by GBP for part of its current traffic to be rerouted so that its overall traffic load is reduced, potentially reducing its power consumption by putting a subset of its physical links to sleep. Each router is aware of the interference-free TH links due to the pre-calculated IRLLs (see Section 5.3.2.2). If the router is the head router of an interference-free Token Holding (TH) link, it becomes a TH router and is responsible for locally offloading that TH link. Since multiple non-interfering TH links can be concurrently selected, GBP can converge quicker compared to other ETE schemes [20][27][40] which are based on purely sequential operations.



**Figure 5-2: Timeline for the online operation of GBP.**

A non-TH router does not offload any logical link unless it becomes the head router of a selected TH link during forthcoming selection rounds. Moreover, a logical link can be selected to become a TH link only once per GBP cycle.

TH routers broadcast an *operation-completed* message when they have finished operating on all their TH links. Routers only broadcast the new TE-LSAs upon receiving *operation-completed* messages from all current TH routers (all routers in the network know which routers are TH routers because they all compute the links which are TH links). Each router then repeats the process of selecting the TH links. The GBP cycle stops after the list of logical links is exhausted, i.e. all logical links have become TH links.

### **5.3.2 Offline Component**

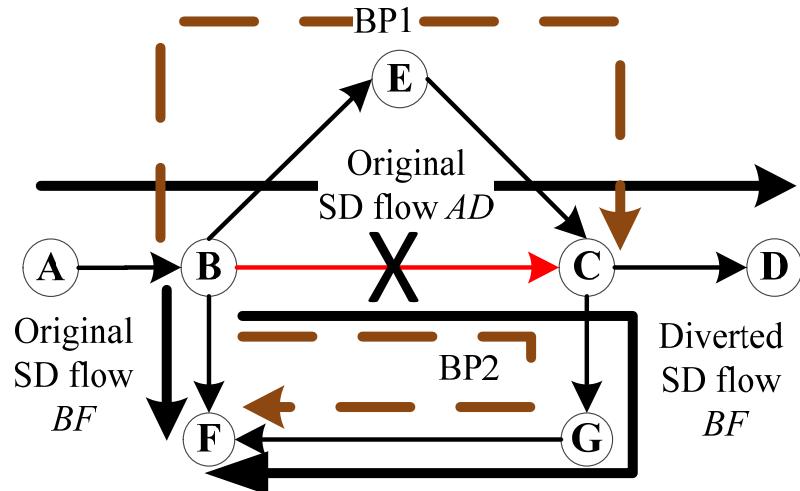
The offline component of GBP consists of two stages. The first stage is responsible for the identification of the backup paths which are eligible for participation in GBP operations and this is done by filtering out the backup paths with excessive end-to-end delay. The second stage is the generation of the IRLL for each logical link of the network. The IRLLs will allow several logical links to concurrently have their traffic diverted without any conflict that could lead to adverse effects on the network such as traffic congestion.

#### **5.3.2.1 Identification of Eligible Backup Paths**

In this stage the offline component of GBP identifies the backup paths that are eligible to participate in GBP operations based on the end-to-end delay of the backup path compared to its protected logical link. Each backup path entry in the MPLS label table of a router has an associated binary bit, *delay\_ok*, which is set to 1 if the path meets the constraint on the maximum path length as determined by the network operator or 0 otherwise. If *delay\_ok* is 0, the backup path is only used for link failure protection. For simplicity, from now on only the identified eligible backup paths are considered.

#### **5.3.2.2 Generation of Interference-Risk Links Lists**

The second stage of the offline component calculates the IRLL for each logical link in the network. The interference-risk links of a TH link are defined as *all logical links which must not be used by other TH links to divert traffic to when that specific TH link is undergoing offloading*. Taking the network topology in Figure 5-3 as example, when logical link  $B \rightarrow C$  becomes a TH link, no other TH links are allowed to divert traffic onto logical link  $B \rightarrow C$  and the logical links of the backup path  $BP1$  which consists of logical links  $B \rightarrow E$  and  $E \rightarrow C$ . Therefore, logical links  $B \rightarrow C$ ,  $B \rightarrow E$  and  $E \rightarrow C$  are in the IRLL of logical link  $B \rightarrow C$ .



**Figure 5-3: Basic network topology to illustrate how IRLLs are generated.**

In addition, a TH link may have some Source-Destination (SD) flows which are currently diverted on it. GBP allows a TH link to divert these flows back to their original respective protected logical links if the head router of the SD flows is also the head router of the TH link. Therefore, logical link  $B \rightarrow F$  is also added to the IRLL of TH link  $B \rightarrow C$  in order to allow  $B \rightarrow C$  to deactivate the backup path  $BP2$  so that the diverted SD flow  $BF$  is re-routed back to the original protected logical link  $B \rightarrow F$ . Hence, TH link  $B \rightarrow C$  has an IRLL consisting of logical links  $B \rightarrow C$  (itself),  $B \rightarrow E$ ,  $E \rightarrow C$  and  $B \rightarrow F$ . The description of how IRLLs are used to avoid interference when multiple TH links are concurrently offloaded will be given in Section 5.3.3.3.

### 5.3.3 Online Component

Figure 5-4 shows a top-level view of the online component of GBP. This component consists of four different stages. At the start of each GBP optimization cycle, each router in the network needs to collect link state information from other routers in the network so as to get an updated and consistent view of the state of the network.

Following this gathering of information, the second stage of GBP is performed where routers may need to deactivate some already activated backup paths because the diversion of traffic on them has led to the logical links constituting these paths to become overloaded. After the deactivation of the overloaded backup paths, GBP continues the optimization process by choosing logical links which have not been offloaded in this optimization cycle and do not conflict with each other according to the IRLLs. Attempts are then made to divert traffic from the selected logical links so that their power consumption and/or utilization go down while not overloading any paths.

The next step is for all routers to broadcast the state of their logical links so that the new state of the network is captured by all routers. In the same manner as before, a new set of unselected logical links is then selected to have their traffic diverted. This iterative process of selecting

logical links to have their traffic diverted is continued until all logical links in the network have been considered in the current GBP optimization cycle. In the remaining part of this section, all the four different stages of the online component of GBP are described in more detail.

### 5.3.3.1 Stage 1: Gathering the State of the Network

At the start of each GBP optimization cycle, all routers need to concurrently collect the broadcasted information about the state of all logical links in the network. This procedure can leverage the TE-LSAs which are already specified in the suite of traffic engineering protocols such as OSPF-TE [74]. GBP requires two types of information about the logical links from the TE-LSAs, namely the current load and value of the *TH\_status\_flag*. Each logical link has a binary bit called *TH\_status\_flag* which is set to 0 at the beginning of each GBP optimization cycle and to 1 after its associated logical link has become TH link in the current GBP optimization cycle. This prevents a logical link from becoming a TH link again in that particular cycle.

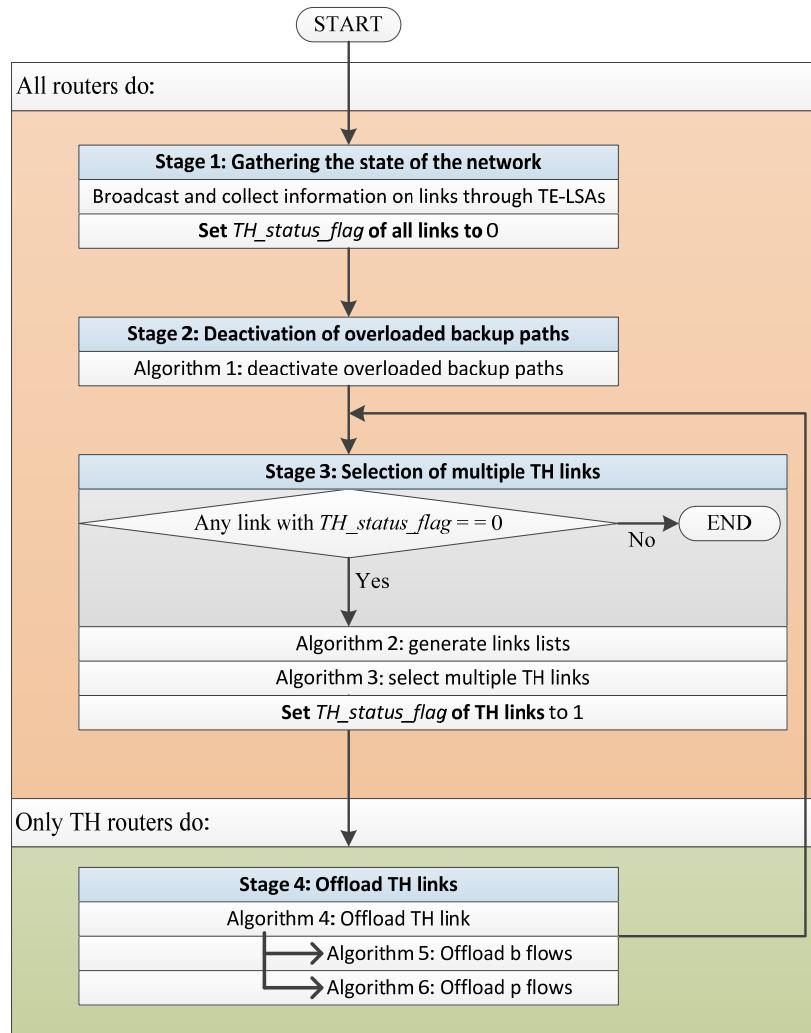


Figure 5-4: Flow chart showing the different operations in one GBP optimization cycle.

### 5.3.3.2 Stage 2: Deactivation of Overloaded Backup Paths

After routers have collected information from TE-LSAs, they verify that the logical links in activated backup paths are not overloaded. This situation may be caused by an increase in traffic volume. If such overloaded backup paths are identified, they are deactivated to relieve their overloaded logical links. The flow chart<sup>5</sup> to demonstrate this operation is given in Figure 5-5 and the algorithmic complexity of *Stage 2* is  $O(|L|^2)$ .

### 5.3.3.3 Stage 3: Selection of Multiple Token Holding Links

The main purpose of *Stage 3* is to calculate which TH links can be selected at the same time without any interference. After the deactivation of the overloaded backup paths, routers broadcast new TE-LSAs so that others are aware of the new state of the network. On receiving the new TE-LSAs, each router forms a list of logical links that excludes all logical links which have their *TH\_status\_flag* equal to 1. Initially all logical links will be included since their respective *TH\_status\_flag* is set to 0 at the start of each GBP optimization cycle. Given that all routers have the same view of the state of the network through the new broadcasted TE-LSAs, they will therefore form an identical list. The list of logical links is partitioned into two disjoint sub-lists, namely *priority\_links\_list* and *normal\_links\_list*. The *priority\_links\_list* contains all the logical links that have their utilization above a predefined threshold  $\alpha$  and therefore, violate the constraint in Eq. (5.5). These have priority to become TH links because if they are successfully offloaded, the resilience of the network against traffic upsurges will improve. The *priority\_links\_list* is then sorted in descending order based on the excess load of the logical links. This excess load,  $\Phi_l$ , is defined by:

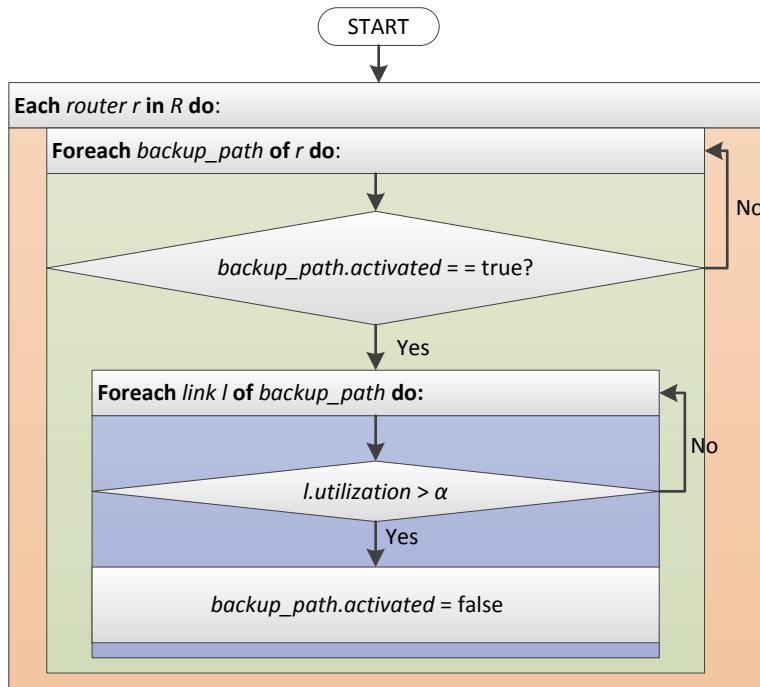
$$\Phi_l = \max \left( \text{mod} \left( \frac{f_l}{z_l} \right), f_l - \frac{\alpha}{100} c_l \right) \quad (5.6)$$

where  $z_l$  is the bandwidth capacity of one physical link of the logical link  $l$ . The first term of the maximum function in Eq. (5.6) represents the excess load on the TH link that prevents the TH link from going to the next lower power level by putting an additional physical link to sleep. The second term is used to calculate the excess load on the TH link that prevents its utilization from dropping below  $\alpha$  % of its total capacity.

The impact of the offloading of highly-utilized logical links on other logical links is minimized by setting a limit on the maximum spare capacity of the other logical links in the network, so that their utilization does not exceed the predefined threshold  $\alpha$ . The design choice of prioritizing the offload of highly-utilized links can be seen as a way of improving the resilience of the network against traffic upsurges.

---

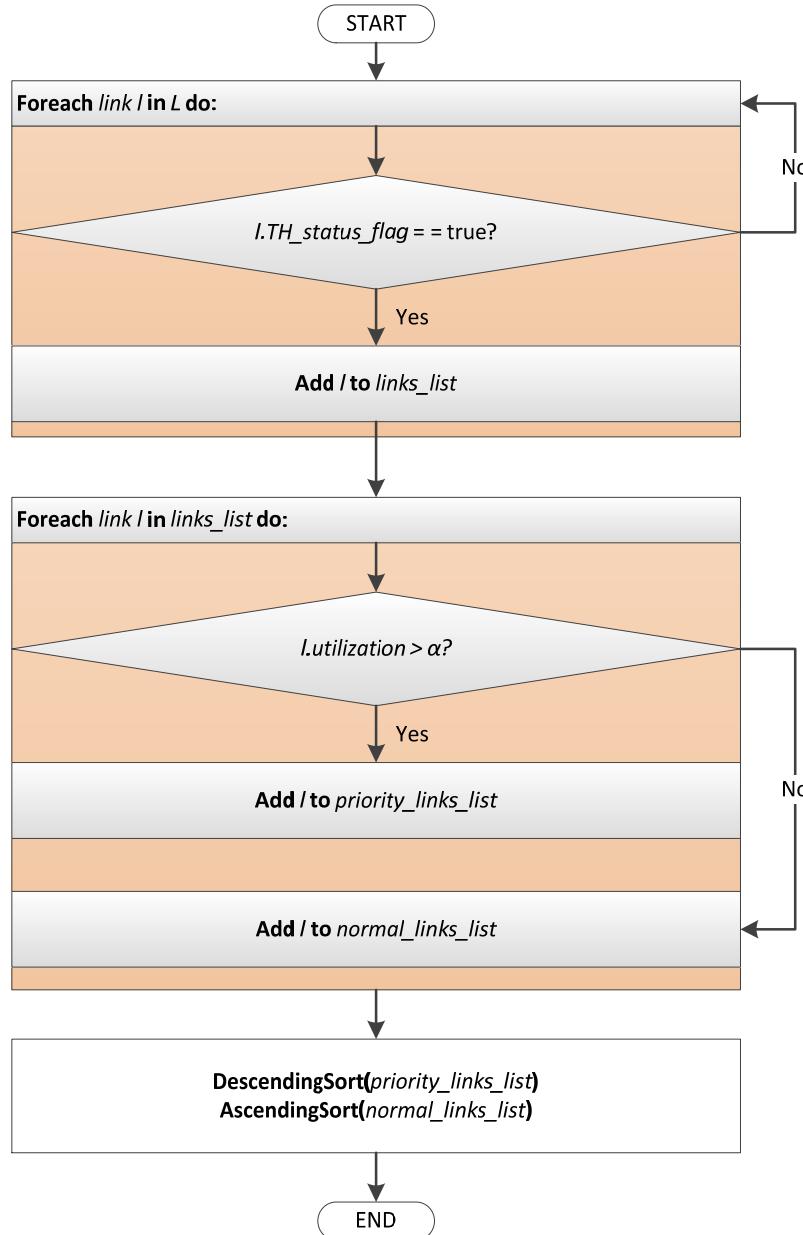
<sup>5</sup> In all flow charts in this chapter,  $y.x$  means  $x$  is a property/variable of  $y$ . In addition,  $y[x]$  means  $y$  is a list where  $x$  is a position in the list.



**Figure 5-5: Flow chart showing the deactivation of overloaded backup paths.**

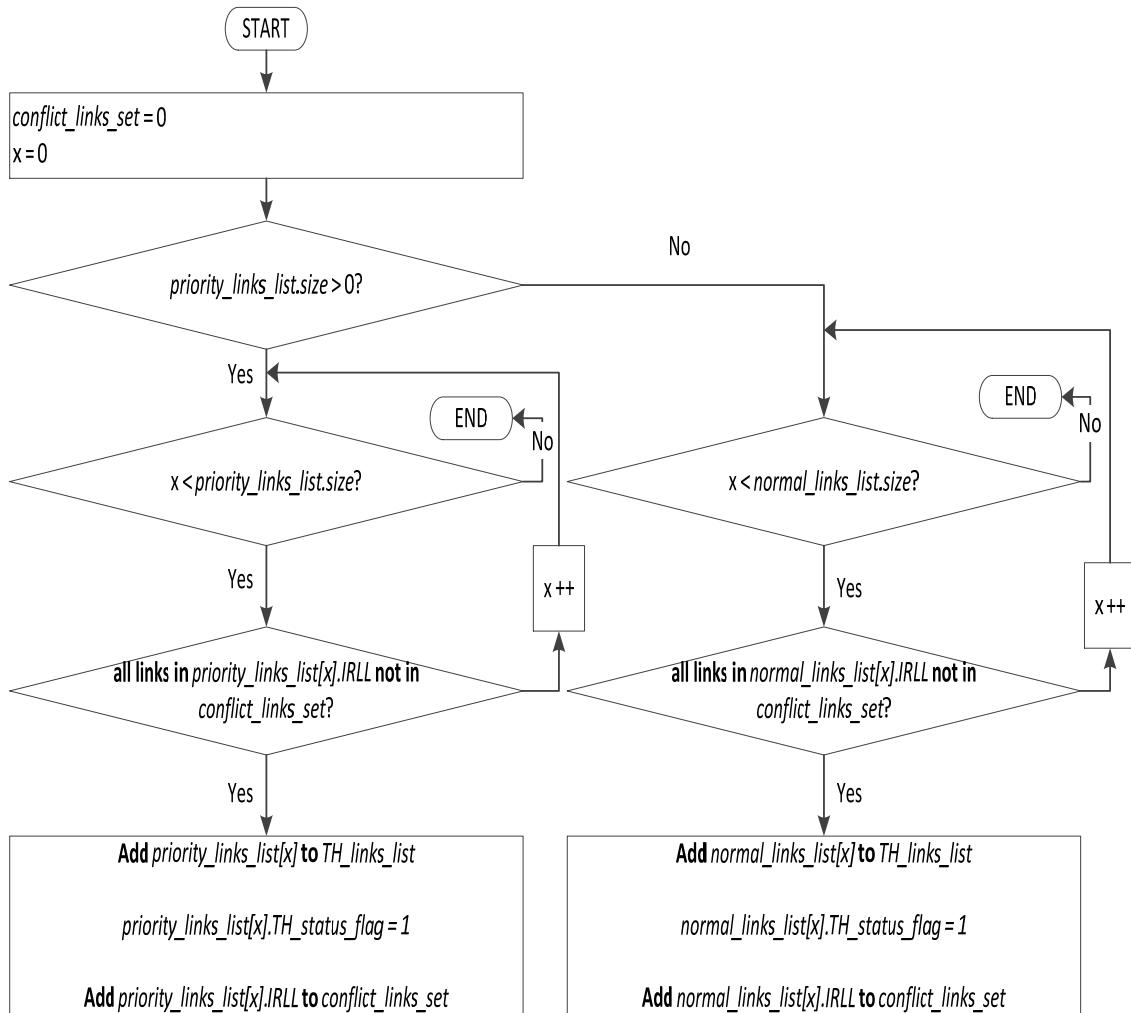
After all logical links from the *priority\_links\_list* have become TH links in the current GBP optimization cycle; it is the turn of the ones in the *normal\_links\_list*. Unlike the previous list, this one is sorted in ascending order according to the excess load, which is calculated by using Eq. (5.6). This is because it is easier to offload small excess loads to alternative paths and therefore, achieve greater power savings. The overall process of generating the two sub-lists is described in Figure 5-6.

The second part of *Stage 3* involves the selection of multiple logical links that can concurrently become TH links. The flow chart for this part is given in Figure 5-7. Each router has an initially empty set called *conflict\_links\_set*, which is also emptied after each iteration of Multiple TH Links Selection in a single GBP optimization cycle. The router goes through the *priority\_links\_list* and if the logical link does not have any logical link of its IRLL in the *conflict\_links\_set*, the router makes the logical link become a TH link and adds its IRLL links to the *conflict\_links\_set*. The example topology in Figure 5-4 can be used to illustrate this process. If the logical link  $B \rightarrow C$  is the first selected TH link in a Multiple TH Links Selection iteration, its IRLL links (i.e., links  $B \rightarrow C$ ,  $B \rightarrow E$ ,  $E \rightarrow C$  and  $B \rightarrow F$ ) are added to the *conflict\_links\_set*. Any subsequent selected TH links in this iteration must have none of their IRLL links in the *conflict\_links\_set*. For example, logical link  $B \rightarrow F$  cannot become TH link in this iteration because some of its IRLL links (i.e.  $B \rightarrow C$  and  $B \rightarrow F$ ) are already in the *conflict\_links\_set*. Since  $B \rightarrow C$  cannot become TH link again during the current GBP optimization cycle,  $B \rightarrow F$  will have the opportunity to become TH link during the next iteration of the Multiple TH Links Selection. When a logical link becomes a TH link, its *TH\_status\_flag* is set to 1.



**Figure 5-6: Flow chart showing the generation of the two sub-lists: *priority\_links\_list* and *normal\_links\_list*.**

After going through the whole *priority\_links\_list*, the router performs the same selection procedure for links in the *normal\_links\_list*. When a router has finished calculating the TH links and under the condition that it is the head router of at least one TH link, it becomes a TH router. A TH router will attempt to offload its TH links through the process in *Stage 4*, described next, and then broadcast an *operation-completed* message to all other routers in the network upon concluding the whole operation. Routers in the network will broadcast a new TE-LSA immediately after they have received the *operation-completed* message from all the current TH routers. The next iterations for the Multiple TH Links Selection in the current GBP optimization cycle can begin after routers receive all the new TE-LSAs. The overall algorithmic complexity of *Stage 3* is  $O(|L|^2)$ .



**Figure 5-7: Flow chart showing how to select multiple TH links.**

When a router has finished calculating the TH links and under the condition that it is the head router of at least one TH link, it becomes a TH router. A TH router will attempt to offload its TH links through the process in *Stage 4*, described next, and then broadcast an *operation-completed* message to all other routers in the network upon concluding the whole operation. Routers in the network will broadcast a new TE-LSA immediately after they have received the *operation-completed* message from all the current TH routers. The next iterations for the Multiple TH Links Selection in the current GBP optimization cycle can begin after routers receive all the new TE-LSAs. The overall algorithmic complexity of *Stage 3* is  $O(|L|^2)$ .

#### 5.3.3.4 Stage 4: Offloading of the Token Holding Link

The overall flow chart for *Stage 4* is given in Figure 5-8. Each TH router has three lists of SD flows for each of its logical links and they are used to classify all the SD flows on the logical link. The first list is the *p\_flows\_list* which is a list of all SD flows that normally use the link, i.e. the flows were not diverted onto the logical link by GBP. The second list is the *b\_flows\_list* which is a list of all SD flows that were diverted onto the logical link by GBP and the logical link has the

same head router as the protected link. The third list is the *s\_flows\_list* which contains all the remaining SD flows on the logical link.

The TH router has direct control over the SD flows in the *p\_flows\_list* and *b\_flows\_list* because the TH router acts as the head router for these flows and therefore, it can decide whether to route these SD flows on either their original protected link or backup path. The selection between routing either on the protected link or backup path is based on where the SD flow is currently routed and whether there is enough spare capacity on the alternate route to support the SD flow without either increasing the power consumption of that route or overloading the logical links of the alternate route. The spare capacity,  $h_l$ , of a logical link  $l$  is given by:

$$h_l = \min \left( z_l - \text{mod} \left( \frac{f_l}{z_l} \right) , \frac{\alpha}{100} c_l - f_l \right) \quad (5.7)$$

where the first term of the minimum function is the amount of traffic that can be added to a logical link without this link going to the next higher power level by waking up an additional physical link. The second term restricts the amount of traffic that can be added to a logical link so that its overall utilization percentage does not go above the predefined threshold  $\alpha$ . If Eq. (5.7) results in less than zero, then  $h_l$  is 0 meaning there is no spare capacity.

In the first step of *Stage 4*, each TH router selects one of its TH links and calculates the excess load on the link using Eq. (5.6). Since the SD flows in the *p\_flows\_list* and *b\_flows\_list* of the TH link are under the direct control of the TH router because it is the head router of these flows, they are the only SD flows targeted for removal from the TH link. The flows in the *b\_flows\_list* of the TH link are first targeted. By rerouting them to their respective original protected links, the TH link will be offloaded and additionally, both delay and wastage of bandwidth will be reduced because of the shorter path taken by the SD flow. The *b\_flows\_list* is sorted in descending order according to load so that the least number of SD flows are moved back to their respective protected link when the excess load on a TH link is removed. Hence, allowing the offloading of the TH link to be quicker because the smallest number of reconfigurations is done.

The decision whether a flow in the *b\_flows\_list* can be moved back to the protected logical link depends on the spare capacity of the respective protected logical link. The spare capacity is calculated by using Eq. (5.7). If the spare capacity of the protected link is larger than the size of the diverted flow, then the flow is added to the list of flows, *flows\_to\_reroute\_list*, to be rerouted. The information contained about the logical links in the TH router needs to be updated to reflect that the traffic in a flow is to be rerouted. As the TH router knows all the logical links involved in the backup path rerouting, it can decrease the load of these logical links and therefore increase their spare capacity.

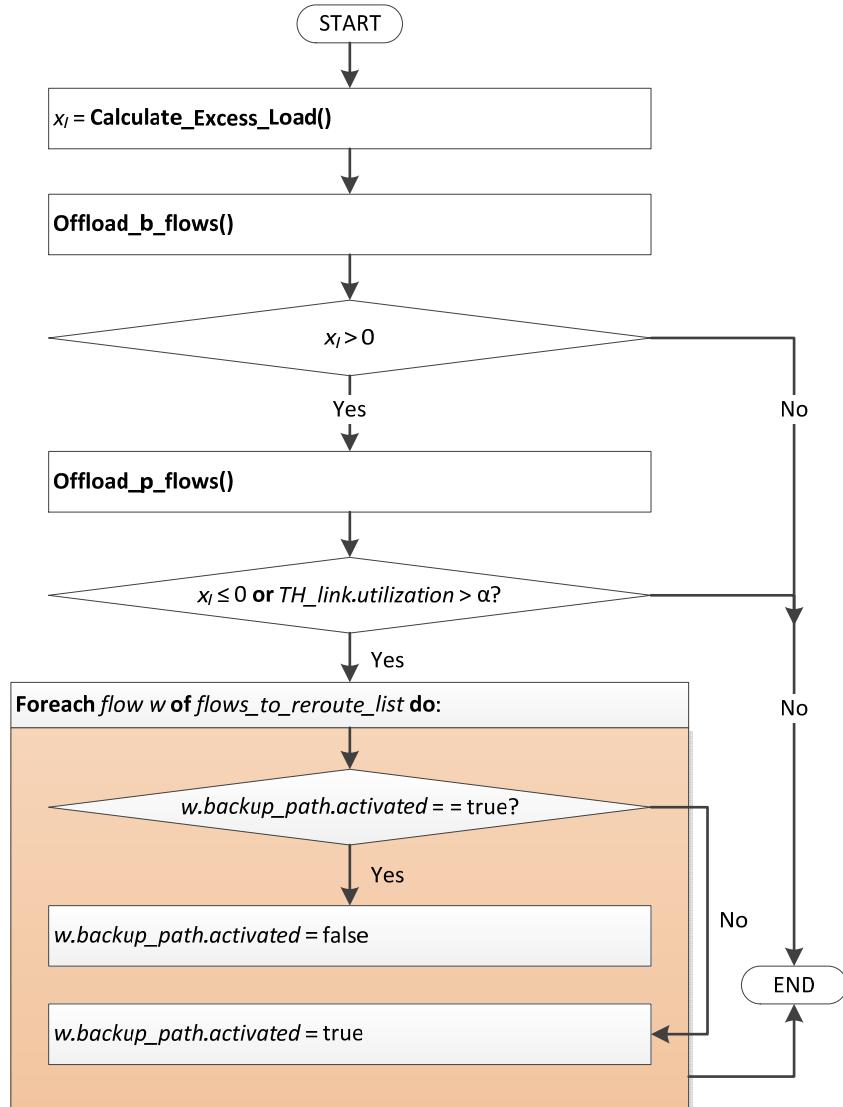
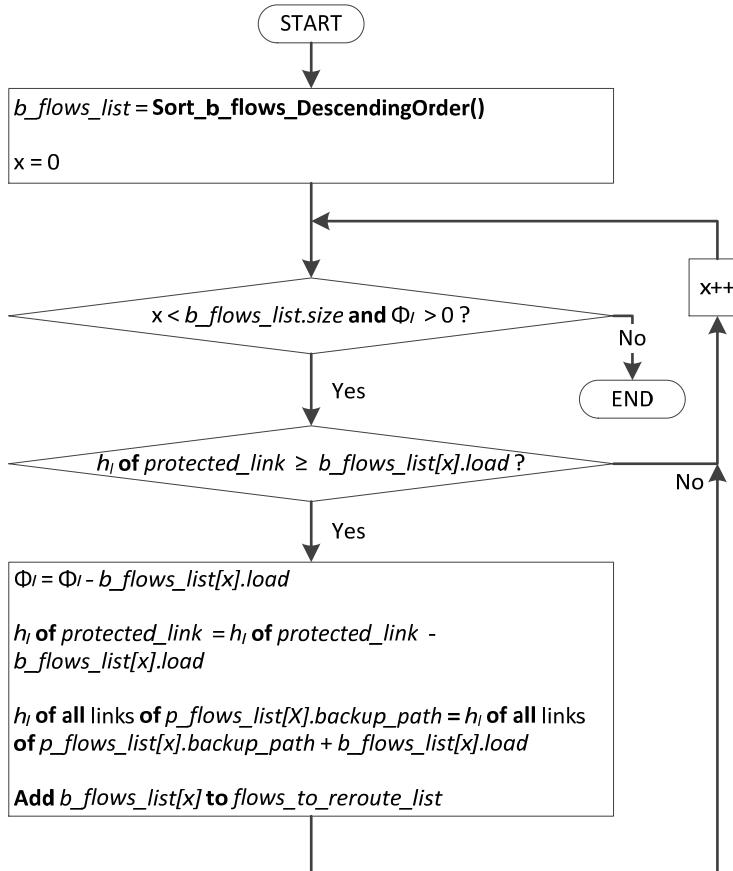


Figure 5-8: Flow chart showing the overall process of offloading a TH link.

The protected logical link of the backup path will have its load increased and consequently its spare capacity decreased because some of its previously-diverted SD flows have been rerouted back on it. The amount of excess load on the TH link is reduced by the total size of the rerouted SD flows. If the excess load is still above zero, the next flow in the *b\_flows\_list* is selected for rerouting. The algorithmic complexity of rerouting diverted SD flows back to their protected link is  $O(|R|^2(|L| + \log|R|^2))$  and the flow chart for this step is presented in Figure 5-9.

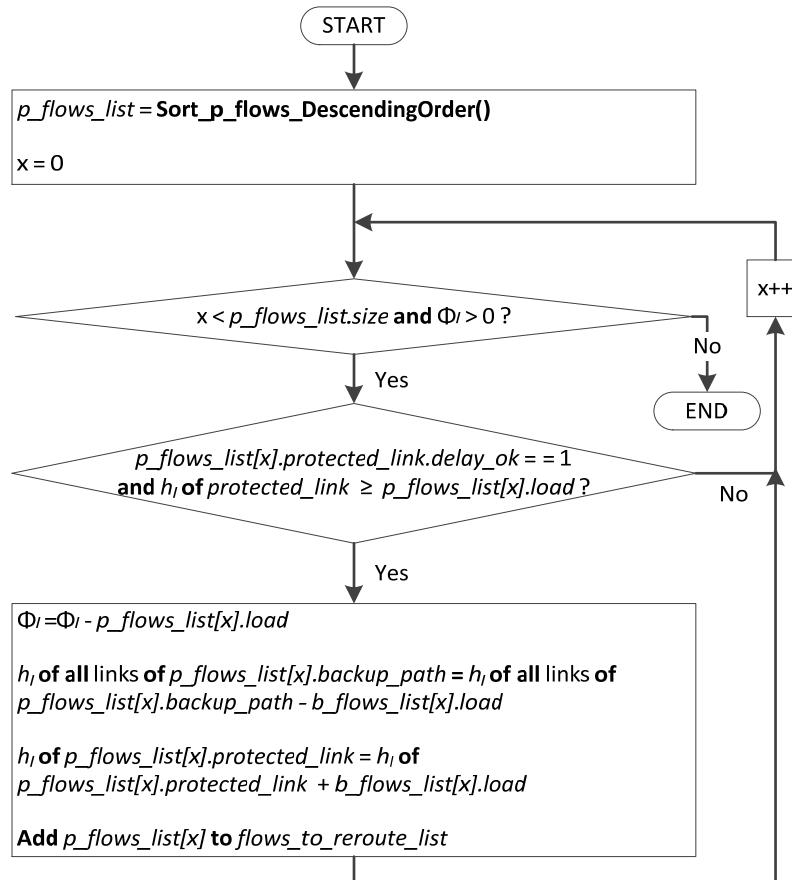
In the third step of *Stage 4* if the excess load is still greater than zero and all the flows in the *b\_flows\_list* have become candidate for rerouting, the flows in the *p\_flows\_list* are then considered for rerouting. The process is similar to the one for flows in the *b\_flows\_list* and its flow chart is presented in Figure 5-10. First, the list is sorted in descending order according to the size of the flows. An SD flow can be diverted to the backup path of the protected TH link if the *delay\_ok* bit of the backup path is equal to 1. This enables the backup path to accept SD flows to be diverted on it.



**Figure 5-9: Flow chart showing the offloading of *b\_flows* from a TH link.**

The spare capacity of the backup path for this SD flow on the TH link is calculated by taking the minimum spare capacity of all the logical links involved in the backup path with the spare capacity of a logical link being calculated using Eq. (5.7). If the spare capacity of the backup path is greater than the size of the SD flow to be rerouted, the SD flow is added to the *flows\_to\_reroute\_list*. The spare capacity and load of the logical links, involved in this flow rerouting, are updated. This process continues until either the excess load is equal/less than zero or all the SD flows in the *p\_flows\_list* have become candidate for rerouting. The algorithmic complexity of offloading the flows is  $O(|R|^2(|L| + \log|R|^2))$ .

The final step of *Stage 4* is to implement all the SD flow reroutes if either the excess load is less/equal to zero or the utilization threshold of the TH link is above the predefined threshold  $\alpha$ . The second criterion is used in the case where the excess load is still greater than zero, meaning that not enough SD flows have been successfully rerouted but it is still desirable to implement all the successful SD flow reroutes because this will decrease the load of an overloaded TH link and make it more resilient to traffic upsurges. Due to the way the spare capacity of a logical link is calculated, it is not possible for GBP to overload a logical link above a set threshold  $\alpha$  while offloading other logical links.



**Figure 5-10: Flow chart showing the offloading of *p\_flows* from a TH link.**

The algorithmic complexity for this step is  $O(|R|^2)$ . The overall algorithmic complexity of *Stage 4* and of a whole GBP optimization cycle are  $O(|R|^2(|L| + \log|R|^2))$  and  $O(|L|^2 + |R|^2(|L| + \log|R|^2))$  respectively.

### 5.3.4 Handling of Logical Link Failures when GBP is active

When GBP is active in a MPLS-enabled backbone network, single logical link failures are handled by two mechanisms: the conventional failure protection mechanism and a GBP enhanced failure-protection mechanism. The conventional failure-protection mechanism is applied regardless of whether GBP is active or not in the network.

When a logical link fails in an MPLS-enabled backbone network, the head router of the failed logical link will divert the SD flows in the failed logical link to its backup path. This conventional failure protection mechanism can be illustrated with the simple example topology in Figure 5-11 where all the logical links have a capacity of 100Mbps. If the logical link  $B \rightarrow C$  fails, its traffic is diverted by the failure protection mechanism onto its backup path  $BP1$ . For example, if  $B \rightarrow C$  was initially carrying 50Mbps of traffic, upon its failure the 50Mbps traffic will be diverted on  $BP1$  which consists of logical links  $B \rightarrow E$  and  $E \rightarrow C$ . It should be noted that the logical links

$B \rightarrow E$  and  $E \rightarrow C$  may be carrying their own traffic (as they can be involved in other default and backup paths), and the diverted traffic from the failed logical link  $B \rightarrow C$  will add to this demand. If the utilization of a logical link is greater than 100% of its capacity, the excess traffic on that link will be lost due to congestion. For example, if  $B \rightarrow E$  and  $E \rightarrow C$  were initially carrying 50Mbps and 60Mbps before the failure of  $B \rightarrow C$ , the utilization of link  $B \rightarrow E$  and  $E \rightarrow C$  will become 100% and 110% after the failure of  $B \rightarrow C$ . Hence, link  $E \rightarrow C$  will suffer from packet loss because its utilization is greater than 100%.

Moreover, GBP incorporates an enhanced failure-protection mechanism which allows it to minimize the probability that any logical link will become over-utilized after single logical link failures. This enhanced failure-protection mechanism has two objectives; the first one is the reduction of the traffic that is diverted from the failed logical link onto its backup path. The rationale behind this objective is to avoid the logical links of the backup path of the failed logical link from becoming over-utilized due to the traffic diversion. The second objective is the increase of the spare capacity of the backup paths because this will allow the backup paths to accommodate diverted traffic without becoming over-utilized.

In order to support this GBP enhanced failure-protection mechanism, the head router of the failed logical link needs to broadcast a failure notification to all routers in the network when the logical link fails. Upon receipt of the failure notification, routers will check and *deactivate* any of their activated backup paths which use any IRLL logical links of the failed logical link. This is done to achieve the two objectives of the GBP enhanced failure-protection mechanism. The deactivation of affected backup paths is done by diverting the traffic on them back onto their protected logical link and as a result, one or more physical links contained by that protected logical link may need to wake up to carry the reverted-back traffic.

In order to illustrate the GBP enhanced failure-protection mechanism, the topology in Figure 5-11 is extended into Figure 5-12 with the same traffic demands still being used and all logical links having capacity of 100Mbps. In this topology, it can be seen that the logical link  $B \rightarrow D$  has part of its traffic diverted onto its backup path  $BP2$  (which uses the failed logical link  $B \rightarrow C$ ) when GBP is active so that additional physical links can go to sleep in  $B \rightarrow D$ .

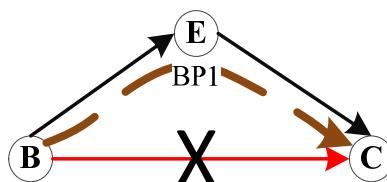


Figure 5-11: Illustrative topology to demonstrate the conventional failure-protection mechanism in an MPLS-enabled network.

Therefore, when  $B \rightarrow C$  fails, the traffic to be diverted is greater compared to the scenario where GBP is not active since  $B \rightarrow C$  is carrying traffic from another protected logical link through the activated  $BP2$ . In this case, the GBP enhanced failure-protection mechanism will deactivate  $BP2$  so as to reduce the traffic to be diverted due to the failure of  $B \rightarrow C$ . For example, if  $BP2$  is diverting 5Mbps on  $B \rightarrow C$  from  $B \rightarrow D$ , the total traffic diverted by  $B \rightarrow C$  onto  $BP1$  when it fails will reduce from 60Mbps to 55Mbps due to the deactivation of  $BP2$ . In order to enable the deactivation of  $BP2$ , it is necessary to deactivate all backup paths that are originally using the protected logical link of  $BP2$ , i.e.  $B \rightarrow D$ . This is done so that  $B \rightarrow D$  has enough spare capacity to accommodate the increased traffic due to the deactivation of  $BP2$ . As mentioned in Section 5.3.2.2,  $B \rightarrow D$  is part of the IRLL of  $B \rightarrow C$  and according to the GBP enhanced failure-protection mechanism; any backup paths which use a link in the IRLL of a failed link need to be deactivated.

Moreover, it may happen that the backup path of a failed logical link has reduced spare capacity because its logical links are part of the backup paths of other logical links. For example in Figure 5-12, if the logical link  $E \rightarrow F$  has part of its traffic diverted onto its backup path  $BP3$  for the power saving operations of GBP, this traffic diversion by GBP will reduce the spare capacity of backup path  $BP1$  of the failed logical link  $B \rightarrow C$ . Therefore,  $BP1$  may become congested when  $B \rightarrow C$  fails. In order to alleviate this problem, it is necessary to deactivate any backup paths that are using any logical links of the backup path of the failed link. For example, if  $BP3$  was initially sending 5Mbps and is deactivated when  $B \rightarrow C$  fails, then the traffic on  $B \rightarrow E$  will fall by 5Mbps to 55Mbps and will have 45Mbps of spare capacity. Since the traffic on  $B \rightarrow C$  is now reduced to 45Mbps due to the previous deactivation of  $BP2$ ,  $BP1$  can now support all the traffic diverted by  $B \rightarrow C$  when it fails.

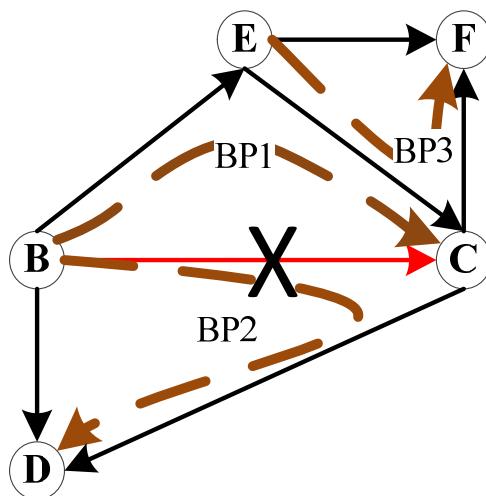


Figure 5-12: Illustrative topology to demonstrate GBP enhanced failure-protection mechanism.

When a logical link fails, the two failure protection mechanisms operate at the same time and independently from one another. If GBP has not previously activated any backup paths which use the IRLL links of the failed logical link, only the conventional failure-protection mechanism will have an effect on the traffic distribution in the network.

## 5.4 GBP Performance Evaluation

In this section, the results of the evaluation of the proposed GBP scheme are presented and discussed. The performance evaluation is done using two operational academic network topologies, as described in Section 5.4.1. More specifically, the following network parameters are measured and discussed: 1) power and energy consumption; 2) MLLU; 3) increase in maximum packet delay; and 4) effect of single logical link failures on the post-failure peak MLLU and energy-efficiency.

### 5.4.1 Network Scenarios

GBP was evaluated by using two academic network topologies, namely GÉANT and Abilene, and their real traffic matrices [4]. The GÉANT topology, summarized in Table 5-3, consists of 23 Points-of-Presence (PoPs) and 74 unidirectional links with different capacities. In Table 5-3,  $|L|$  represents the number of logical links of a specific capacity  $c$  that have  $y_l$  number of physical links that individually transmit at  $\lambda$  optical carrier speed and consume  $p$  amount of power. The power consumption for each physical link was obtained from the maximum power consumption of Cisco line cards [70]. For a physical link of capacity of OC-48, a one-port line card uses 140W and for OC-3, there is no one-port line card but rather a four-port line card with total power consumption of 196W. For simplicity, an OC-3 physical link is assumed to consume 196/4=49W. The Abilene topology consists of 12 PoPs and 30 unidirectional links of varying capacity, as shown in Table 5-4 (which have the same notation as in Table 5-3).

For the traffic demands in the GÉANT and Abilene network, 480 consecutive traffic matrices that were measured at 15-minute intervals were considered [4]. Consistently, the 15-minute interval was also adopted as the period of the GBP optimization cycle. That is, the application of each traffic matrix on the network corresponds to the starting point in time of a new optimization cycle of GBP.

### 5.4.2 Power and Energy Saving Gains

The power saved by GBP was calculated by using Eq. (5.8) below where  $y_l^o$  is the number of physical links which are sleeping in the logical link  $l$ . In order to evaluate the power saving gains

of GBP, it was compared with a Theoretical Upper Bound (TUB) scheme. TUB was obtained with IBM CPLEX [75] by adding the concept of link sleeping to the conventional non-integer Multi-Commodity Flow problem. Therefore, the restriction of using only the predefined protected links and backup paths to route traffic demands is not applied in TUB. GBP was simulated with different values of  $\alpha$  for both GÉANT and Abilene.

$$\text{Power Saved} = \frac{\sum_{l=1}^{|L|} y_l^\rho \times p_l}{\sum_{l=1}^{|L|} y_l \times p_l} \times 100\% \quad (5.7)$$

Figure 5-13 and Figure 5-14 show that for all traffic matrices, GBP was able to save a significant amount of power for both simulated networks. The values of power savings in both figures represent the power saved as a percentage of the total power consumed by the networks when all links are active. Understandably, there is a gap between GBP and TUB in terms of power saving performance because GBP uses a single path for each SD flow while TUB uses a large number of paths. Of course, it should be noted that the path configuration given by TUB cannot be implemented in practice because of the large number of paths between each SD pair in the network that would be required.

**Table 5-3: Power model of GÉANT network topology.**

Logical Links				Physical Links			
L	c (Mbps)	y <sub>l</sub>	L  × y <sub>l</sub>	z (Mbps)	λ	p (W)	L  × y <sub>l</sub> × p (W)
32	9953	4	128	2488	OC-48	140	17920
2	4876	2	4	2488	OC-48	140	560
32	2488	1	32	2488	OC-48	140	4480
8	155.2	1	8	155.2	OC-3	49	392
<b>Σ</b>	<b>74</b>		<b>172</b>				<b>23352</b>

**Table 5-4: Power model of Abilene network topology.**

Logical Links				Physical Links			
L	c (Mbps)	y <sub>l</sub>	L  × y <sub>l</sub>	z (Mbps)	λ	p (W)	L  × y <sub>l</sub> × p (W)
28	9920	4	112	2480	OC-48	140	15680
2	2480	1	2	2480	OC-48	140	280
<b>Σ</b>	<b>30</b>		<b>114</b>				<b>15960</b>

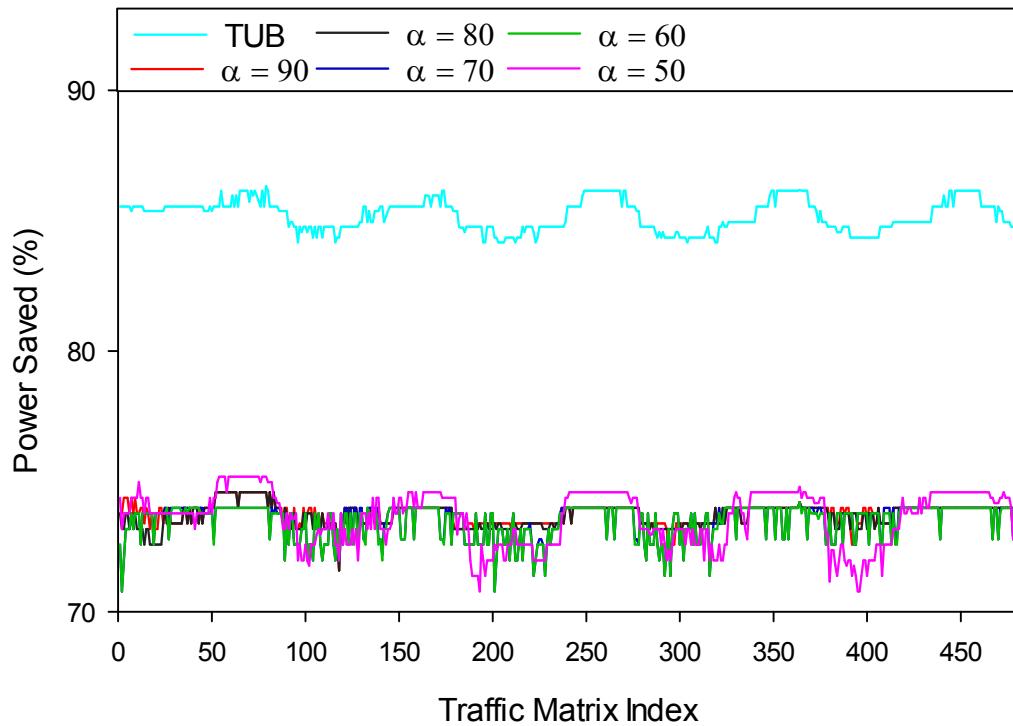


Figure 5-13: Power saved for the GÉANT topology.

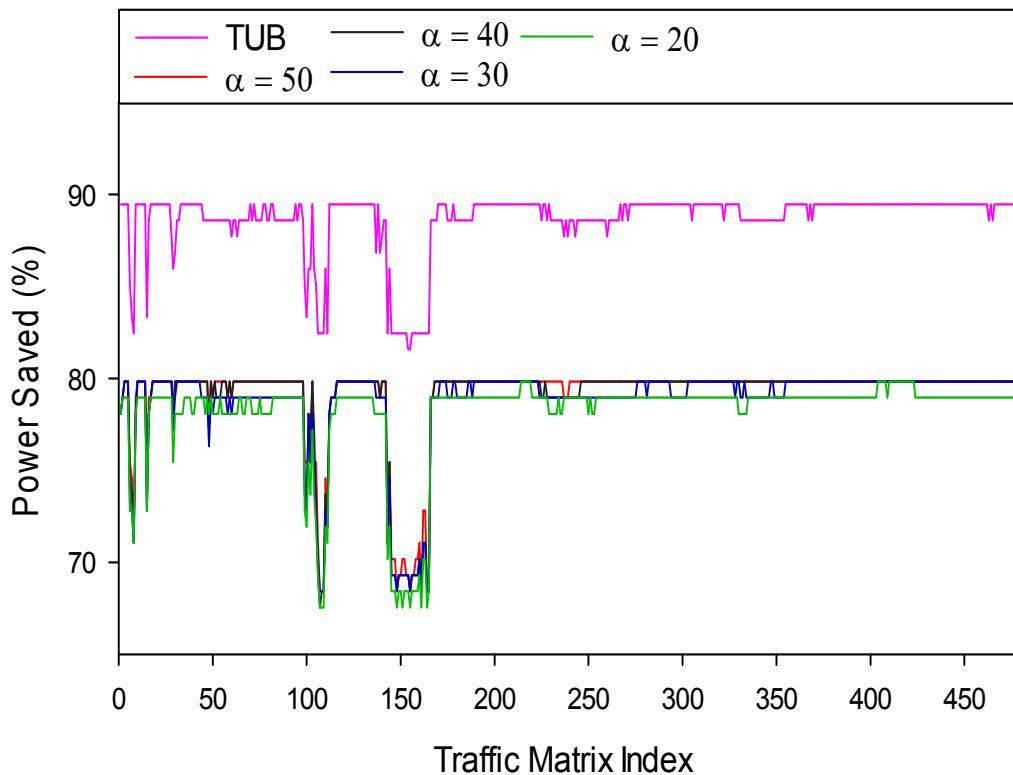


Figure 5-14: Power saved for the Abilene topology.

Figure 5-13 shows that, for the GÉANT scenario, power savings do not change significantly when  $\alpha$  is reduced from 90 to 50. The energy-efficiency,  $E$ , of GBP is calculated by using Eq. (5.9) which calculated the total energy saved by GBP as a proportion of the theoretical optimal energy that could be saved . The value of  $\Psi$  does not change much when  $\alpha$  is reduced in Table 5-5, which correlates with the observation made for Figure 5-13. The observations made in Figure 5-13 and Table 5-5 can be explained by the fact that even when  $\alpha$  is high, GBP does not have a high degree of freedom to divert a significant amount of traffic onto a backup path because this will make the logical links constituting that backup path to consume a larger amount of power. That is, the spare capacity of most backup paths in the network remains mostly constant when  $\alpha$  is varied from 90 to 50 because the first term in the spare capacity equation, Eq. (5.7), is the dominating one for most logical links in the network.

$$E = \frac{\text{Energy Saved by GBP}}{\text{Theoretical Optimal Energy Saved}} \times 100 \% \quad (5.9)$$

Figure 5-14 and Table 5-5 show that the power saved for the Abilene network also does not change much with the variation of  $\alpha$ . It is interesting to see from the performance curves in Figure 5-14 that GBP reacts to sudden changes in traffic conditions as TUB does even though the paths available to “absorb” these changes are limited for GBP.

### 5.4.3 Maximum Logical Link Utilization

The dynamicity of the Maximum Logical Link Utilization (MLLU) resulting from the GBP operations across the evaluation period is shown in Figure 5-15 and Figure 5-16 for GÉANT and Abilene respectively. The original MLLU values that were measured in the network are also included in the figures. For GÉANT, GBP was able to offload highly-utilized logical links and the maximum MLLU became close to the value of  $\alpha$  when this was varied from 90 to 70. This shows that GBP is able to successfully enforce the constraint in Eq. (5.5) by offloading over-utilized logical links while not overloading under-utilized ones as a result of its operations.

When  $\alpha$  is further reduced from 70 to 50, GBP is unable to reduce the peak MLLU to meet the value of  $\alpha$  because there are not enough logical links with sufficiently large spare capacity to carry the traffic from the logical links whose utilization is above  $\alpha$ . During its operations, GBP does not divert traffic on any logical link whose utilization is greater than  $\alpha$  and therefore, it does not make the peak MLLU becomes worse compared to when GBP is not operated. Moreover, GBP will not divert an excessive amount of traffic to any logical links because this may result in the utilization of the logical links going above  $\alpha$ . GBP can enforce these two restrictions on the diversion of traffic by calculating the spare capacity of a logical link through Eq. (5.7).

**Table 5-5: Energy-efficiency,  $E$ , for the GÉANT and Abilene topology.**

GÉANT		Abilene	
$\alpha$	$E$ (%)	$\alpha$	$E$ (%)
90	86.6	50	89.2
80	86.5	40	89.1
70	86.2	30	88.8
60	86.2	20	88.1
50	86.4		

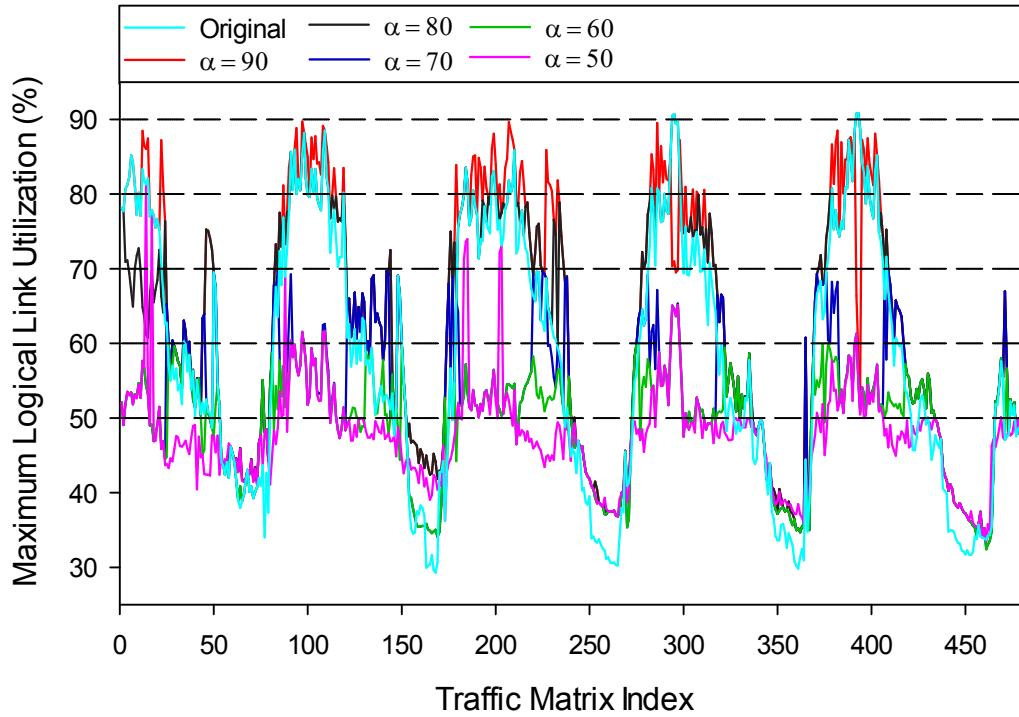
Figure 5-16 shows that the MLLU experiences substantial and frequent fluctuations during the original operation of Abilene where GBP was not activated. This is also reflected in the change in MLLU during the operation of GBP. The peak MLLU remains the same as the original one when  $\alpha$  is varied from 50 to 20 during the operation of GBP. This happens when there is not enough space capacity on the backup paths to accept diverted traffic because they are already carrying a high volume of traffic. Hence, GBP cannot reduce the peak MLLU for Abilene as it did for GÉANT due to a lack of spare capacity in the Abilene network.

For some traffic matrices in Figure 5-15 and Figure 5-16, it can be observed that the MLLU values of GBP can go up when compared to the original ones because GBP wants to concentrate traffic on the minimum number of logical links possible so as to save the maximum amount of power. This concentration of traffic on a minimum number of logical links is always done with the constraint that the MLLU should not go above the value of  $\alpha$  because of the rerouting actions of GBP. For the traffic matrices for which the MLLU is above  $\alpha$  for GBP, the original MLLU is also above  $\alpha$  even though GBP is not being operated. This shows that GBP is not responsible for breaking the constraint  $\alpha$ , but it is just the original high volume of traffic that causes the MLLU to be above  $\alpha$ .

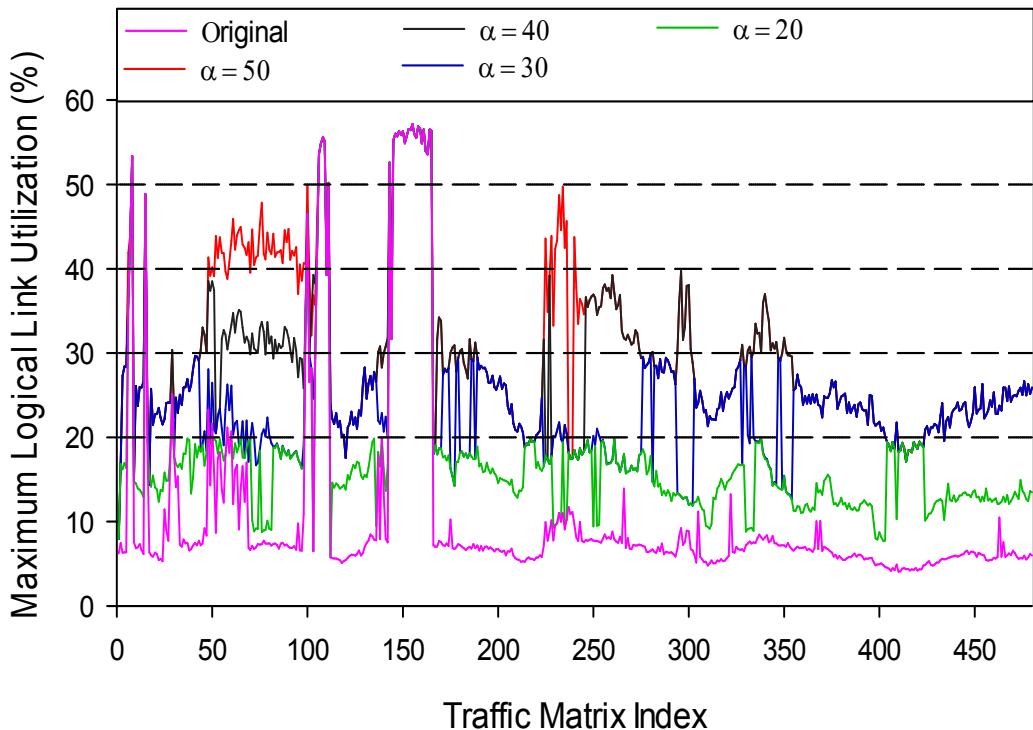
#### 5.4.4 Increase in Maximum Packet Delay

Table 5-6 shows the increase in the maximum packet delay when GBP is operated for GÉANT and Abilene respectively. The maximum packet delay was measured in the same way as described in Section 3.5.2. For the GÉANT network topology, the average increase in the maximum packet delay is small, at most 6.39ms. For the maximum increase in the maximum packet delay, the increase was not higher than 21.9ms. For the Abilene network topology, the average and maximum increase in maximum packet delay were quite small at around 1.58ms and 6.96ms respectively. There was no change in the observed minimum maximum packet delay for both network scenarios. The main conclusion from these delay results is that it can be assumed that GBP does not significantly affect the

packet delay in a well-connected network, which is the case of GÉANT and Abilene. This is despite the fact that GBP reroutes some traffic on longer backup paths to offload logical links.



**Figure 5-15: Variation of MLLU for Original and GBP for GÉANT topology.**



**Figure 5-16: Variation of MLLU for Original and GBP for Abilene topology.**

**Table 5-6: Increase in maximum packet delay for the GÉANT and Abilene topology.**

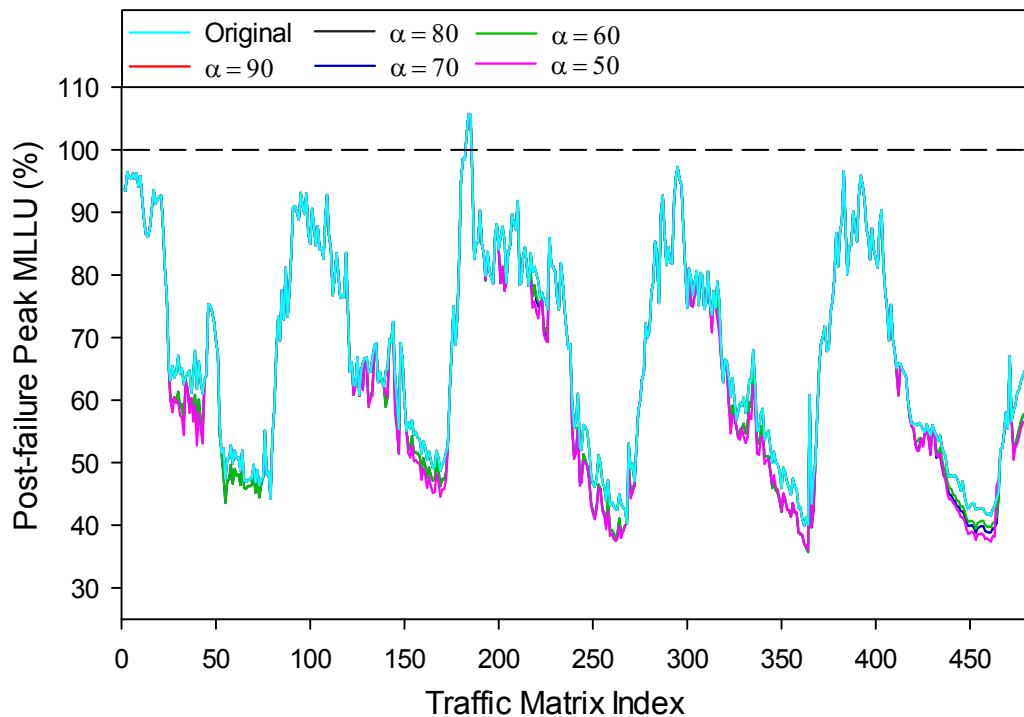
GÉANT				Abilene			
A	Increase in Maximum Packet Delay (ms)			$\alpha$	Increase in Maximum Packet Delay (ms)		
	Max.	Avg.	Min.		Max.	Avg.	Min.
90	17.5	5.72	0	50	6.96	1.44	0
80	20.8	6.17	0	40	6.96	1.58	0
70	13.8	6.38	0	30	6.96	1.53	0
60	13.8	6.39	0	20	6.96	1.52	0
50	21.9	5.26	0				

### 5.4.5 Single Logical Link Failure Analysis

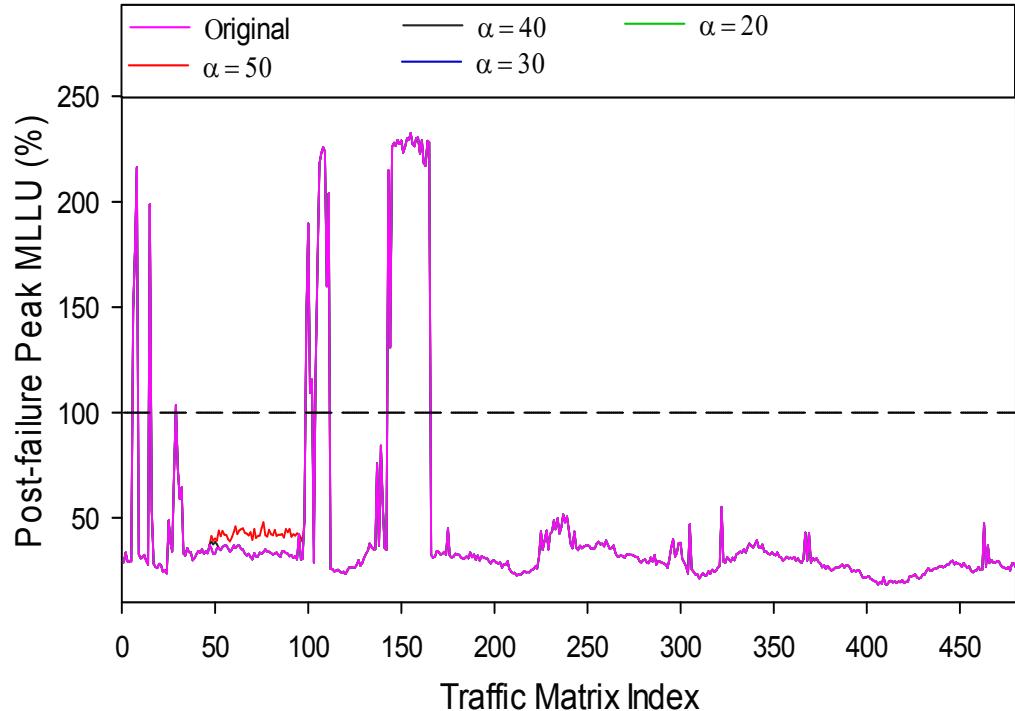
A whole logical link can fail due to the cut of all its bundled physical fiber links, e.g. due to earthquakes or other physical damages. Since GBP makes use of backup paths for energy savings in addition to their primary purpose of preventing packet loss upon single logical link failures, it is important to investigate the effect of GBP on the post-failure peak MLLU [76] and the energy savings upon any single logical link failure. As mentioned in Section 5.3.4, upon the failure of an active logical link, the failure protection mechanism will divert traffic from that failed logical link to its corresponding backup path. This failure protection mechanism of logical links is applied in an MPLS-enabled network regardless of whether GBP is in place or not. In addition, GBP has an enhanced failure-protection mechanism which enables the amount of traffic to be diverted from the failed logical link to be reduced and also for the spare capacity of the backup path of the failed link to be increased so that the backup path can support the traffic diverted from the failed logical link.

#### 5.4.5.1 Post-failure Peak Maximum Logical Link Utilization

Single logical link failures were simulated for all considered traffic matrices of both GÉANT and Abilene scenarios. The aim was to examine the effect of the dual use of the resources of the backup paths for both power savings and logical link failure protection. The effect of GBP on logical link failure protection was quantified by computing the peak MLLU after any logical link has failed in the network. Figure 5-17 and Figure 5-18 show the post-failure peak MLLU for normal energy-agnostic operation (i.e. GBP is not operated) and when GBP is active with different values of  $\alpha$  for GÉANT and Abilene respectively. It can be observed that during single logical link failures, there was no increase in post-failure peak MLLU when GBP is active in most cases. When the traffic matrices for which the post-failure peak MLLU value is higher than 100% were further analyzed in detail, it could be observed that the peak MLLU values is the same irrespective of whether GBP is active or not.



**Figure 5-17:** Post-failure peak MLLU between no-GBP and GBP operations for the GÉANT topology.



**Figure 5-18:** Post-failure peak MLLU between no-GBP and GBP operations for the Abilene topology.

This is an important observation since it is only when the peak MLLU is higher than 100% that traffic is actually lost in the network and therefore, the same peak MLLU suggests that the network suffers from the same degree of over-utilization for these traffic matrices for both the scenarios where GBP is activated or not. For the peak MLLU values which are less than 100%, it can be observed in Figure 5-17 that the post-failure peak MLLU can sometimes be reduced when GBP is active for GÉANT. This can happen because GBP can reduce the MLLU in the network as shown in Figure 5-15 for GÉANT. Therefore, there is a probability that the post-failure peak MLLU will be lower when GBP is activated because there is more spare capacity on the backup paths to accommodate the diverted traffic during single logical link failures. For Abilene in Figure 5-18, there are a few traffic matrices where the post-failure peak MLLU is greater under GBP when the peak MLLU values are less than 100%. This is because GBP concentrates traffic on the lowest number of logical links as possible so that the physical links in the other logical links can go to sleep. Therefore, there may not be enough spare capacity on some protected logical links in the network when the GBP enhanced failure-protection mechanism diverts traffic to them, when it deactivates their respective backup paths because these backup paths use any IRLL links of the failed logical link.

The main overall conclusion from Figure 5-17 and Figure 5-18 is that backup paths can be used for greater power savings without reducing the ability of the backup paths to prevent packet loss during single logical link failures. Therefore, there is no need to provision additional paths which are dedicated for power savings because non-conflicting use of the backup paths for power savings and prevention of packet loss during single link failures is feasible.

#### **5.4.5.2 Impact on Energy Saving Gains during Single Logical Link Failures**

The energy saving gains by GBP can be affected by three different factors when a logical link fails in the network. The first factor is that a failed logical link will not consume any energy, i.e. all the physical links of the failed logical links are considered to be “sleeping” and not consuming power. On the other hand, it may be necessary for logical links involved in the activated backup path of a failed logical link to wake up additional physical links in order to provide the extra spare capacity required to accommodate the diverted traffic from the failed logical link without causing any post-failure traffic congestion. This is the second factor which can affect the energy saving gains. In addition, some already activated backup paths (for energy savings and/or reduction of utilization at their protected logical links) need to be deactivated so as to reduce the amount of traffic to be diverted from the failed logical links, and also to increase the spare capacity of the backup path of the failed logical links. Therefore, the third factor is the possible reduction in the power consumption of the logical links involved in the deactivated backup paths while the protected logical links of the deactivated paths may consume more power because of the

increased traffic on them. In GBP, the head routers are responsible for determining how many sleeping physical links in each of their logical links should wake up, so that only the minimum number of physical links is active without causing any traffic congestion. This design choice is made so as to maximize the energy savings in the network but without compromising on post-failure packet loss.

Table 5-7 shows the change in energy-efficiency during single logical link failures compared to a failure-free scenario for GÉANT and Abilene respectively. In this table, a positive number means logical link failures increase the energy saved in the network compared to a failure-free scenario while a negative number means the opposite. On average, the energy saved decreases during single logical link failures. This is because several physical links will have to wake up in the logical links, constituting the backup path of the failed logical link, if there is not enough spare capacity in the currently active physical links of the backup path. The power consumption of the protected logical links of the deactivated backup paths also increases due to increased traffic on them. However, the maximum reduction in energy saved during single logical link failures is not significant.

Interestingly, the Table 5-7 indicates some unexpected observations where the energy saving gains can further increase upon single link failures. This can be explained by the fact that logical link failures are similar to putting links to sleep but with the key difference that the MLLU constraint in Eq. (5.5) needs not be respected during single logical link failures. In other words, GBP cannot always achieve the same level of energy savings compared to when some single logical links fail because GBP would break the MLLU constraint defined in Eq. (5.5). However, single logical link failures do not have this restriction because during these events, logical links can be loaded with as much traffic as the failure protection mechanism can handle.

**Table 5-7: Change in energy-efficiency,  $\Delta E$ , for the GÉANT and Abilene topology.**

GÉANT				Abilene			
$\alpha$	$\Delta E$ (%)			$\alpha$	$\Delta E$ (%)		
	Max.	Avg.	Min.		Max.	Avg.	Min.
90	-0.0000847	-3.91	-12.0	50	0.0970	-1.37	-2.61
80	-0.0752	-3.89	-11.9	40	0.153	-1.33	-2.65
70	-0.345	-3.59	-8.20	30	0.476	-1.19	-2.70
60	-0.344	-3.58	-8.20	20	0.987	-0.834	-3.13
50	0.0365	-3.61	-8.66				

## 5.5 Summary

There are backbone networks, such as Abilene [4], which do not exhibit a regular traffic pattern and it is important to develop a suitable online ETE scheme which can significantly improve the energy-efficiency of such networks while preserving the ability of network operators to do load-balancing. In order to meet these requirements, a new computationally-efficient scheme called *Green Backup Paths* (GBP) was designed.

GBP improves the energy-efficiency of a network by leveraging existing MPLS backup paths which were primarily installed for the failure protection of logical links. This dual use of backup paths leads to a lower management overhead of paths in the network. It has been demonstrated in the evaluation section that GBP does not adversely affect the ability of backup paths to prevent packet loss during single logical link failures. Furthermore, GBP is a fully-distributed scheme where multiple logical links can be concurrently offloaded so that their number of active physical links reduces. This results in the reduced power consumption of the whole network.

The ability of GBP to save power was evaluated through simulations on two academic networks, namely GÉANT and Abilene, and their respective real traffic matrices. Results showed that GBP achieves significant power savings while also decreasing the maximum logical link utilization and maintaining the maximum packet delay in the network within reasonable bounds. The power savings achieved by GBP were always within 15% of the Theoretical Upper Bound which indicates the efficiency of GBP.

The design of GBP may seem complex at first view but this is the required level of complexity that is necessary to obtain the best energy-savings and load-balancing performance possible when the network has dynamic traffic demands. The extensive network monitoring system is required by GBP in order to make accurate decisions, which is made much more difficult when the size of the traffic demands is changing quickly. The network monitoring system is already available on current network equipment nowadays and network operators only have to activate it and make the monitoring system broadcasts frequently the state of the network. The previous two schemes designed in this thesis, TLS and GLA, assumed that the traffic demands will not deviate much from the historical traffic demands but this is not guaranteed and therefore, TLS and GLA may use the simple method of traffic scaling to account for this phenomenon. In addition, GBP is aware that the link between any two routers can actually be made of a bundle of physical links and therefore, it can leverage this feature of the network to achieve greater energy-efficiency compared with the first two schemes designed in this thesis, TLS and GLA,

# Chapter 6

## 6 Conclusions

### 6.1 Thesis Summary

The power consumption of backbone networks is of increasing concern to network operators because their electricity bill is rapidly increasing as they deploy more network devices to meet the traffic demands of their users. To make matters worse, the price of the unit of electricity has been increasing consistently in the past decades. Hence, network operators are looking for novel ETE schemes that can drastically decrease the power consumption of the network so as to significantly reduce their electricity bills. While network operators desires reduced power consumption, they still need to preserve targeted network performances such as low maximum link utilization, small maximum packet delay, high stability of new network configurations and high resilience to packet loss during single link failures. This thesis presented three novel ETE schemes which attempt to reduce the overall power consumption of backbone networks but without deteriorating the original network performances.

Three ETE schemes have been designed in this thesis because not all backbone networks exhibit the same traffic pattern and therefore, custom ETE schemes must be designed to make the most out of each particular traffic pattern. The first ETE scheme designed is TLS which targets backbone networks that exhibit a diurnal traffic pattern. The assumption of a regular diurnal traffic allows a time-driven approach to be adopted for the reconfiguration of the network. In order to avoid instabilities in the network due to frequent reconfigurations, only two network configurations are used for each day. Furthermore, the start and end time of each configuration is the same for each day so as to ease the implementation of TLS. The ability of TLS to save a significant amount of energy was evaluated on the PoP representation of the GÉANT academic network and its real traffic matrices. TLS was able to obtain up to 28% energy savings without exceeding the MLU performance during peak time. Moreover, it was demonstrated that it is possible to use only the traffic matrices of one week only to calculate an energy-efficient off-peak network configuration that is valid for several weeks. Hence, there is no need for frequent recalculations of the network configuration.

Single link failures can happen frequently in a backbone network and therefore, TLS was extended into TLS-SLFP to cater for these network events. The main objective of TLS-SLFP is to

improve the resilience of the off-peak network configuration of TLS against single link failures so that the network does not suffer from packet loss during these network events. Since improved resilience against single link failures requires more spare capacity to be present in the network, the off-peak network configuration of TLS-SLFP must have more active links. The larger number of active links results in the energy-efficiency of TLS-SLFP to be lower than for plain TLS, for e.g. TLS-SLP achieves 16% energy savings compared to 28% for plain TLS for the same network conditions in the GÉANT European academic network.

The second ETE scheme designed in this thesis is GLA. The primary aim of GLA is to optimize the IGP link weights in a network so that other existing ETE schemes based on link sleeping can achieve further enhanced the energy-efficiency, load-balancing and maximum packet delays. In other terms, GLA configures backbone networks so that other existing ETE schemes can achieve a better performance. As far as is known, this is the first time that link weights are optimized for the purpose of jointly improving the energy-efficiency and load-balancing in operational backbone networks. GLA was evaluated by using three existing ETE schemes, including the TLS scheme. In the first part of the evaluation of GLA, it was shown that GLA can significantly improve the energy-efficiency given by these schemes while achieving close to optimal load-balancing. Unfortunately, the chosen ETE schemes do not take into account the maximum packet delays when optimizing the number of and specific links to put to sleep in order to save energy. Therefore, the chosen ETE schemes can cause an excessive increase in maximum packet delay with as end result a reduction in the QoS. In order to solve this problem, the inner workings of the chosen ETE schemes were modified so that they do not increase the maximum packet delay and therefore, become delay-aware. The evaluations of the new delay-aware ETE schemes showed that significant energy savings can still be achieved while keeping near-optimal load-balancing and experiencing no increase in the maximum packet delay. The comparison between the new delay-aware ETE schemes and their plain counterparts showed that there is a reduction in energy savings for the delay-aware schemes but this is to be expected since a trade-off is being made between energy savings and maximum packet delay.

The third and last ETE scheme designed in this thesis is GBP. In contrast with the first and second ETE schemes, GBP is an online and distributed ETE scheme which can react rapidly to sudden and large changes in traffic demands. GBP leverages pre-installed backup paths to improve the energy-efficiency of MPLS-enabled backbone networks. These backup paths were installed in MPLS-enabled backbone networks so that traffic can be diverted on them during single link failures and therefore, preventing packet loss during these network events. The links of the backup paths are already active since they carry their own traffic and the diversion of traffic on the backup paths does not cause the power consumption of these links to increase. Simulations of single link failures in the two academic networks GÉANT and Abilene have shown that it is

possible to use these backup paths for energy savings without reducing the resilience of the network against packet loss during single link failures. The ability of GBP to make concurrent and interference-free diversions of traffic shows that GBP is a truly-distributed algorithm. The distributed nature of GBP allows it to converge quicker and to scale to large networks. GBP was evaluated through the use of the PoP representation of two academic networks, namely GÉANT and Abilene, and their real traffic matrices. While GÉANT exhibits a regular diurnal pattern for most of the time, Abilene can experience sudden and large fluctuations in its traffic demands. GBP was able to achieve significant energy savings, which are quite close to the theoretical upper bound, while not increasing the maximum logical link utilization and suffering only from a small increase in maximum packet delay.

The first two ETE schemes designed in this thesis, TLS and GLA, are simpler compared to GBP because they operate in an offline fashion and do not have to deal with potentially sudden and unpredictable changes in traffic demands. Therefore, TLS and GLA is only suitable for networks which exhibit a regular and predictable traffic pattern while GBP can operate in networks which have dynamic traffic demands since it is an online ETE scheme. The main advantage with TLS and GLA is that they only use existing features of current network equipment and there is no need to make any changes to these equipment for these schemes to work properly. Therefore, this involves less cost for network operators. In practice, network operators who wish to implement an offline ETE scheme should use the MPD-aware version of TLS together with GLA-SH. The MPD-aware version of TLS- will give network operator the ability to control the maximum packet delay according to its policy and GLA-SH will provide the optimized IGP link weights which allow TLS to achieve significant energy-savings. Depending on the network policy, the network operator may also wish to implement the SLFP version of TLS which will give the network protection against packet loss during single link failures. All the versions of TLS offer the network operator the ability to control the maximum link utilization, at the expense of energy savings, according to its policy.

In order for GBP to tolerate these dynamic traffic behaviors, it has to extend the capability of current network equipment by giving them the ability to make changes in the routing of traffic based on the information received from the network monitoring system. In practice, only a firmware extension for the routers is required so that the routers are able to process the TE-LSAs broadcasted and make routing decisions which will improve the energy-efficiency and load-balancing performance of the network. The network monitoring system that GBP requires is already present in current network equipment and is used to gather information about traffic demands for network planning purposes. It is this network monitoring system which provided the historical traffic demands that the offline ETE schemes TLS and GLA used in order to calculate new network configurations.

In the future, network operators will deploy more energy-efficient hardware and network protocols and since ETE neither changes the amount of traffic in the network nor changes the characteristic of the power consumption of network devices, ETE is complementary to these green initiatives.

## 6.2 Thesis Conclusions

The main conclusions that can be drawn from this thesis are:

- 1) TLS has shown that it is possible to design an offline ETE scheme which can achieve significant energy savings with only two network configurations and therefore, there is no need to reconfigure the network frequently to achieve significant energy savings since this will cause instabilities in the network;
- 2) TLS-SLFP has shown that it is possible for the reduced topologies calculated by ETE schemes to provide a similar amount of resilience to traffic congestion during single link failures as full topologies while still achieving significant energy savings;
- 3) GLA has demonstrated that the load-balancing and energy-efficiency of many existing offline ETE schemes can be improved significantly by optimizing the IGP link weights in backbone networks. This is because the original link weights are only optimized when the full topology is used and are no longer optimized when links are put to sleep;
- 4) It has been shown that ETE schemes can excessively increase the maximum packet delay in backbone networks and that ETE schemes should take into account their effect on the maximum packet delay during their operation. All ETE schemes designed in this thesis ensures that the maximum packet delay remains within a predetermined bound set by the network operator during their operation;
- 5) GBP has shown that it is possible to re-use existing backup paths for energy savings while not preventing the backup paths from performing their primary function, i.e. avoid packet loss during single link failures. Therefore, there is no need to install dedicated paths for energy savings in the network with as end result a lower management overhead of paths; and
- 6) A light-weight conflict-avoidance mechanism can be designed to allow the different decision entities of an online ETE scheme to concurrently make conflict-free decisions.

### 6.3 Future Work

In the future, all the ETE schemes that have been designed in this thesis will need to be implemented in a test bed in order to validate the behavior that was observed during simulations. This validation will give network operators a greater level of confidence about the effectiveness of the different ETE schemes.

All the newly-designed ETE schemes in this thesis consider the traffic demands to be inelastic but this may not be the case in an operational environment. This is because it is well-known that a large amount of the traffic demand in a network is actually for the same content from multiple users [77][78]. Therefore, current research efforts in fields such as Content Delivery Networks (CDNs) and Information Centric Networks (ICNs) can reduce the traffic in a network and enable the newly-designed ETE schemes to reduce the energy consumption of the network further because the spare capacity in the network will increase. Furthermore, network operators may strategically place several caching servers in their network so as to further reduce the traffic in their network without having to rely on the traditional CDN providers.

Currently, the power consumption of network devices does not vary much with the amount of traffic being forwarded. This is likely to change in the near future because manufacturers and researchers are investing a lot of effort in making the relationship between the power consumption and traffic forwarded by network devices become as close as possible to linear. The new power profiles of network devices will lead researchers to design novel ETE schemes which will make the most out of these new power profiles. Researchers in [8] have already done an initial evaluation of different possible future power profiles. If the power profiles of network devices become convex, then it will be possible to design a fully-distributed ETE scheme which can achieve global optimality in negligible convergence time.

In this thesis, the newly-designed ETE schemes are targeted at backbone networks only. It would be interesting to do a holistic analysis of the power consumption of the whole wired network, which spans the backbone, aggregation and edge segments of the network. From this holistic analysis, it would be possible to design an ETE scheme for each segment of the network so that they can cooperate with one another and reduce the overall power consumption of the whole wired network. In addition, it would be interesting to do a cost-benefit analysis of putting caching servers as close as possible to the edge of the network.

# Bibliography

- [1] R. Bolla, F. Davoli, R. Bruschi, K. Christensen, F. Cucchietti, and S. Singh, “The potential impact of green technologies in next-generation wireline networks: Is there room for energy saving optimization?” *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 80–86, August 2011.
- [2] S. Nedevschi, L. Popa, G. Iannaccone, S. Ratnasamy, and D. Wetherall, “Reducing network energy consumption via sleeping and rate-adaptation,” in *Proc. 2008 of the 5th USENIX Symposium on Networked Systems Design and Implementation*, 2008, pp. 323–336.
- [3] C. Lange, “Energy-related aspects in backbone networks,” in *Proc. 2009 of 35th European Conference on Optical Communication*, (Wien, AU), September 2009.
- [4] “The totem traffic engineering toolbox.” Available online: <http://totem.run.montefiore.ulg.ac.be>
- [5] M. Roughan, A. Greenberg, C. Kalmanek, M. Rumsewicz, J. Yates, and Y. Zhang, “Experience in measuring backbone traffic variability: models, metrics, measurements and meaning,” in *Proc. 2002 of ACM SIGCOMM Workshop on Internet Measurement*, 2002, pp. 91–92.
- [6] A. Nucci, A. Sridharan, and N. Taft, “The problem of synthetically generating IP traffic matrices: initial recommendations,” *SIGCOMM Comput. Commun. Rev.*, vol. 35, no. 3, pp. 19–32, July 2005.
- [7] J. Sommers, P. Barford, and B. Eriksson, “On the prevalence and characteristics of MPLS deployments in the open internet,” in *Proc. 2011 of Internet Measurement Conference*, 2011, pp. 445–462.
- [8] J. Restrepo, C. Gruber, and C. Machuca, “Energy Profile Aware Routing,” in *Proc. 2009 of ICC Communications Workshops*, 2009, pp. 1–5.
- [9] J. G. Koomey, “Estimating total power consumption by servers in the U.S. and the world,” Lawrence Berkley National Laboratory, Tech. Rep., February 2007.
- [10] G. Shen and R. Tucker, “Energy-Minimized Design for IP Over WDM Networks,” *IEEE J. Opt. Commun. Netw.*, vol. 1, no. 1, pp. 176–186, June 2009.
- [11] M. Zhang, C. Yi, B. Liu, and B. Zhang, “GreenTE: Power-Aware Traffic Engineering,” in *Proc. 2010 of IEEE International Conference on Network Protocols*, October 2010.

- [12] M. M. Luca Chiaraviglio, Dejan Ciullo and M. Meo, “Modeling Sleep Modes Gains with Random Graphs,” in *Proc. 2011 of INFOCOM Workshop on Green Communications and Networking*, Shanghai, China, April 10–15, 2011.
- [13] R. Hay, D. Chalupsky, E. Mann, J. Tsai, and A. Wertheimer, “Active/Idle Toggling with 0BASE-x for Energy Efficient Ethernet,” IEEE 802.3az Task Force, November 2007. Available online: [http://www.ieee802.org/3/az/public/nov07/hays\\_1\\_1107.pdf](http://www.ieee802.org/3/az/public/nov07/hays_1_1107.pdf)
- [14] “Greentouch green meter research study: Reducing the net energy consumption in communications networks by up to 90% by 2020,” GreenTouch Consortium, Tech. Rep., June 2013.
- [15] X. Dong, A. Lawey, T. El-Gorashi, and J. Elmirghani, “Energy-efficient core networks,” in *Proc. 2012 of Conference on Optical Network Design and Modeling (ONDM)*, pp. 1–9.
- [16] C. Gunaratne, K. Christensen, B. Nordman, and S. Suen, “Reducing the energy consumption of ethernet with adaptive link rate,” *IEEE Transactions on Computers*, vol. 57, pp. 448–461, 2008.
- [17] L. Chiaraviglio, M. Mellia, and F. Neri, “Reducing power consumption in backbone networks,” in *Proc. 2009 of the IEEE International Conference on Communications*, June 2009, pp. 2298–2303.
- [18] B. Puype, W. Vereecken, D. Colle, M. Pickavet, and P. Demeester, “Power reduction techniques in multilayer traffic engineering,” in *Proc. 2009 of Conference on Transparent Optical Networks*, July 2009, pp. 1–4.
- [19] A. Cianfrani, M. Listanti, V. Eramo, M. Marazza, and E. Vittorini, “An energy saving routing algorithm for a green OSPF protocol,” in *Proc. 2010 of INFOCOM*, 2010.
- [20] W. Fisher, M. Suchara, and J. Rexford, “Greening backbone networks: reducing energy consumption by shutting off cables in bundled links,” in *Proc. 2010 of ACM SIGCOMM workshop on Green networking*, 2010, pp. 29–34.
- [21] A. A. Kist and A. Aldraho, “Dynamic topologies for sustainable and energy efficient traffic routing,” *Comput. Netw.*, vol. 55, pp. 2271–2288, June 2011.
- [22] S. Singh and C. Yiu, “Putting the cart before the horse: merging traffic for energy conservation,” *Communications Magazine, IEEE*, vol. 49, no. 6, pp. 78–82, June 2011.
- [23] C. Hu, C. Wu, W. Xiong, B. Wang, J. Wu, and M. Jiang, “On the design of green reconfigurable router toward energy efficient internet,” *Communications Magazine, IEEE*, vol. 49, no. 6, pp. 83–87, June 2011.

- [24] A. C. F. Cuomo, A. Abbagnale and M. Polverini, “Keeping the Connectivity and Saving the Energy in the Internet,” in *Proc. 2011 of INFOCOM 2011 Workshop on Green Communications and Networking*, Shanghai, China, April 10–15, 2011.
- [25] G. N. Rosario G. Garropo, Stefano Giordano and M. Pagano, “Energy aware routing based on energy characterization of devices: Solutions and analysis,” in *Proc. 2011 of 4th International Workshop on Green Communications*, Kyoto, Japan, June 2011.
- [26] A. Aldraho and A. Kist, “Heuristics for dynamic topologies to reduce power consumption of networks,” in *Proc. 2010 of Telecommunication Networks and Applications Conference (ATNAC)*, 2010, pp. 31–36.
- [27] B. Mumey, J. Tang, and S. Hashimoto, “Enabling green networking with a power down approach,” in *Proc. 2012 of IEEE International Conference on Communications*, June 2012, pp. 2867–2871.
- [28] E. Amaldi, A. Capone, and L. Gianoli, “Energy-aware IP traffic engineering with shortest path routing,” *Computer Networks*, vol. 57, no. 6, pp. 1503–1517, 2013.
- [29] F. Giroire, D. Mazauric, J. Moulierac, and B. Onfroy, “Minimizing routing energy consumption: From theoretical to practical results,” in *Proceedings of the 2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing*, 2010, pp. 252–259.
- [30] M. Gupta and S. Singh, “Greening of the internet,” in *Proc. 2003 of SIGCOMM*, 2003, pp. 19–26.
- [31] C. Gunaratne, K. Christensen, and B. Nordman, “Managing energy consumption costs in desktop PCs and LAN switches with proxying, split TCP connections, and scaling of link speed,” *Int. J. Netw. Manag.*, vol. 15, pp. 297–310, September 2005.
- [32] N. Vasic and D. Kostic, “Energy-aware traffic engineering,” in *Proc. 2010 of International Conference on Energy-Efficient Computing and Networking*, 2010, pp. 169–178.
- [33] E. Gelenbe and C. Morfopoulou, “A Framework for Energy-Aware Routing in Packet Networks,” *The Computer Journal*, 2010.
- [34] C. Panarello, A. Lombardo, G. Schembra, L. Chiaraviglio, and M. Mellia, “Energy saving and network performance: a trade-off approach,” in *Proc. 2010 of International Conference on Energy-Efficient Computing and Networking*, 2010, pp. 41–50.
- [35] K. H. Ho and C. C. Cheung, “Green distributed routing protocol for sleep coordination in wired core networks,” in *Proc. 2010 of 6th International Conference on Networked Computing*, May 2010, pp. 1–6.

- [36] A. C. Orgerie, L. Lefevre, and I. Guerin-Lassous, “Energy-efficient bandwidth reservation for bulk data transfers in dedicated wired networks,” *The Journal of Supercomputing*, pp. 1–28, 2011.
- [37] R. Bolla, R. Bruschi, A. Cianfrani, and M. Listanti, “Enabling backbone networks to sleep,” *IEEE Network*, vol. 25, no. 2, pp. 26–31, March-April 2011.
- [38] N. Vasic, D. Novakovic, S. Shehar, P. Bhurat, M. Canini, and D. Kostic, “Identifying and Using Energy-Critical Paths,” in *Proc. 2011 of ACM CoNEXT*, December 2011.
- [39] A. Bianzino, L. Chiaravaggio, and M. Mellia, “Grida: A green distributed algorithm for backbone networks,” in *Proc. 2011 of IEEE GreenCom*, September 2011, pp. 113–119.
- [40] M. Charalambides, D. Tuncer, L. Mamatas, and G. Pavlou, “Energy-Aware Adaptive Network Resource Management,” in *Proc. 2013 of the IFIP/IEEE Integrated Management Symposium*, May 2013.
- [41] M. Shen, H. Liu, K. Xu, N. Wang, and Y. Zhong, “Routing on demand: toward the energy-aware traffic engineering with OSPF,” in *Proc. 2012 of the 11th International IFIP TC6 Conference on Networking*, 2012, pp. 232–246.
- [42] Y. Zhao, S. Wang, S. Xu, X. Wang, X. Gao, and C. Qiao, “Load balance vs energy efficiency in traffic engineering: A game theoretical perspective,” in *Proc. 2013 of INFOCOM*, pp. 530–534.
- [43] G. Athanasiou, K. Tsagkaris, P. Vlacheas, D. Karvounas, and P. Demestichas, “Multi-objective traffic engineering for future networks,” *Communications Letters, IEEE*, vol. 16, no. 1, pp. 101–103, 2012.
- [44] A. Coiro, F. Iervini, and M. Listanti, “Distributed and adaptive interface switch off for internet energy saving,” in *In Proc. of 2011 of International Conference on Computer Communications and Networks (ICCCN)*, August 2011, pp. 1–8.
- [45] E. Gelenbe and T. Mahmoodi, “Distributed energy-aware routing protocol,” in *Computer and Information Sciences II*, E. Gelenbe, R. Lent, and G. Sakellari, Eds. Springer London, 2012, pp. 149–154.
- [46] L. Chiaravaggio, M. Mellia, and F. Neri, “Minimizing ISP network energy cost: formulation and solutions,” *Networking, IEEE/ACM Transactions on*, vol. 20, no. 2, pp. 463–476, 2012.
- [47] E. Palkopoulou, D. Schupke, and T. Bauschert, “Energy efficiency and capex minimization for backbone network planning: Is there a tradeoff?” in *Proc. 2009 of Advanced Networks and Telecommunication Systems (ANTS)*, December 2009, pp. 1–3.

- [48] X. Dong, T. El-Gorashi, and J. Elmirghani, “On the Energy Efficiency of Physical Topology Design for IP Over WDM Networks,” *Lightwave Technology, Journal of*, vol. 30, no. 12, pp. 1931–1942, 2012.
- [49] A. Cianfrani, V. Eramo, M. Listanti, M. Polverini, and A. Vasilakos, “An OSPF-integrated routing strategy for QoS-aware energy saving in IP backbone networks,” *Network and Service Management, IEEE Transactions on*, no. 99, pp. 1–14, 2012.
- [50] S. S. Lee, P.-K. Tseng, and A. Chen, “Link weight assignment and loop-free routing table update for link state routing protocols in energy-aware internet,” *Future Generation Computer Systems*, vol. 28, no. 2, pp. 437–445, 2012.
- [51] E. Amaldi, A. Capone, L. Gianoli, and L. Mascetti, “A MILP-based heuristic for energy-aware traffic engineering with shortest path routing,” in *Network Optimization*, ser. Lecture Notes in Computer Science, Eds. Springer Berlin Heidelberg, 2011, vol. 6701, pp. 464–477.
- [52] B. Fortz and M. Thorup, “Increasing Internet Capacity Using Local Search,” *Comput. Optim. Appl.*, vol. 29, pp. 13–48, October 2004.
- [53] R. Bolla, R. Bruschi, F. Davoli, and A. Ranieri, “Performance constrained power consumption optimization in distributed network equipment,” in *Proc. 2009 of ICC Communications Workshops*, June 2009, pp. 1–6.
- [54] A. Wierman, L. Andrew, and A. Tang, “Power-aware speed scaling in processor sharing systems,” in *Proc. 2009 of INFOCOM*, April 2009, pp. 2007–2015.
- [55] R. Bolla, R. Bruschi, A. Carrega, and F. Davoli, “Green networking with packet processing engines: Modeling and optimization,” vol. PP, no. 99, 2013, pp. 1–1.
- [56] K. W. Kwong, R. Guerin, A. Shaikh, and S. Tao, “Balancing performance, robustness and flexibility in routing systems,” *Network and Service Management, IEEE Transactions on*, vol. 7, no. 3, pp. 186 –199, September 2010.
- [57] A. Bianzino, C. Chaudet, D. Rossi, and J. Rougier, “A survey of green networking research,” *IEEE Commun. Surveys Tutorials*, no. 99, pp. 1–18, 2010.
- [58] S. Uhlig, B. Quoitin, J. Lepropre, and S. Balon, “Providing public intradomain traffic matrices to the research community,” *SIGCOMM Comput. Commun. Rev.*, vol. 36, pp. 83–86, January 2006.
- [59] P. Psenak, S. Mirtorabi, A. Roy, L. Nguyen, and P. Pillay-Esnault, “Multi-Topology (MT) Routing in OSPF,” RFC 4915 (Proposed Standard), Internet Engineering Task Force, June 2007. Available online: <http://www.ietf.org/rfc/rfc4915.txt>

- [60] T. Przygienda, N. Shen, and N. Sheth, “M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs),” RFC 5120 (Proposed Standard), Internet Engineering Task Force, February 2008. Available online: <http://www.ietf.org/rfc/rfc5120.txt>
- [61] “Yin Zhang’s Abilene traffic matrices.” Available online: <http://www.cs.utexas.edu/users/-yzhang/research/AbileneTM/>
- [62] B. Y. Choi, S. Moon, Z. L. Zhang, K. Papagiannaki, and C. Diot, “Analysis of point-to-point packet delay in an operational network,” *Comput. Netw.*, vol. 51, no. 13, pp. 3812–3827, September 2007.
- [63] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, and C. Diot, “Measurement and analysis of single-hop delay on an IP backbone network,” *IEEE J.Sel. A. Commun.*, vol. 21, no. 6, pp. 908–921, September 2006.
- [64] N. Wang, K. Ho, G. Pavlou, and M. Howarth, “An overview of routing optimization for internet traffic engineering,” *IEEE Commun. Surveys Tutorials*, vol. 10, no. 1, pp. 36–56, 2008.
- [65] L. Chiaraviglio, M. Mellia, and F. Neri, “Energy-aware backbone networks: A case study,” in *Proc. 2009 of the IEEE International Conference on Communications*, June 2009, pp. 1–5.
- [66] F. Francois, N. Wang, K. Moessner, and S. Georgoulas, “Optimization for Time-driven Link Sleeping Reconfigurations in ISP Backbone Networks,” in *Proc. 2012 of IEEE Network Operations and Management Symposium*, April 2012.
- [67] F. Francois, N. Wang, K. Moessner, and S. Georgoulas, “Optimizing Link Sleeping Reconfigurations in ISP Networks with Off-Peak Time Failure Protection,” *Network and Service Management, IEEE Transactions on*, vol. 10, no. 2, pp. 176–188, June 2013.
- [68] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, April 2002.
- [69] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti, “Energy Efficiency in the Future Internet: A Survey of Existing Approaches and Trends in Energy-Aware Fixed Network Infrastructures,” *IEEE Commun. Surveys Tutorials*, no. 99, pp. 1–22, 2010.
- [70] “Cisco 12000 series packet over sonet/sdh (pos) line cards.” Available online: [http://www.cisco.com/en/US/prod/collateral/routers/ps6342/product\\_data\\_sheet0900aecd803fd7b9.pdf](http://www.cisco.com/en/US/prod/collateral/routers/ps6342/product_data_sheet0900aecd803fd7b9.pdf)

- [71] F. Hoffmeister and T. Back, “Genetic algorithms and evolution strategies: Similarities and differences,” in *Parallel Problem Solving from Nature*, ser. Lecture Notes in Computer Science, Eds. Springer Berlin Heidelberg, 1991, vol. 496, pp. 455–469.
- [72] A. Raj and O. C. Ibe, “A survey of IP and multiprotocol label switching fast reroute schemes,” *Computer Networks*, vol. 51, no. 8, pp. 1882–1907, 2007.
- [73] H. Alnuweiri, L. Wong, and T. Al-Khasib, “Performance of new link state advertisement mechanisms in routing protocols with traffic engineering extensions,” *IEEE Commun. Mag.*, vol. 42, no. 5, pp. 151–62, may 2004.
- [74] K. Kompella and Y. Rekhter, “OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS),” RFC 4203 (Proposed Standard), Internet Engineering Task Force, October 2005.
- [75] “IBM ILOG CPLEX Optimizer,” September 2013. Available online: <http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/>
- [76] K. H. Ho, N. Wang, G. Pavlou, and C. Botsaris, “Optimizing post-failure network performance for IP fast reroute using tunnels,” in *Proc. 2008 of the 5th ICST Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, 2008, pp. 441–447.
- [77] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, and B. Ohlman, “A survey of information-centric networking,” *Communications Magazine, IEEE*, vol. 50, no. 7, pp. 26–36, 2012.
- [78] J. Choi, J. Han, E. Cho, T. Kwon, and Y. Choi, “A survey on content-oriented networking for efficient content delivery,” *Communications Magazine, IEEE*, vol. 49, no. 3, pp. 121–127, 2011.