Connection Routing and Configuration in Optical Burst Switching Networks

Chen Qian

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

September, 2008

Acknowledgement

First and foremost, I would like to take this opportunity to express sincere gratitude to my supervisor, Associate Professor Mohan Gurusamy, and co-supervisor, Professor Chua Kee Chaing for all the support throughout my PhD candidature. This thesis would not have existed without their guidance and inspiration. Their fruitful discussions with me were instrumental in shaping my research attitude and outlook.

I would also like to thank all the members of Optical Network Engineering (ONE) lab who have made it an enjoyable place to work. And I would also like to thank the lab officer, Mr. David Koh, for his kind support.

I am especially grateful to my parents and husband for their endless love and encouragement. They are my incessant source of hope and happiness throughout my ups and downs.

Abstract

Optical burst switching (OBS) is a promising technology to transfer bursty traffic over wavelength division multiplexed (WDM) networks. As the optical buffers are very expensive and they provide very short delays only, the core nodes in OBS networks are usually bufferless. We identify and analyze the unique features that arise from the bufferless property and consider these features to design efficient schemes to route and configure connections. We assume that the network has Multiple Protocol Label Switching (MPLS) control and the bursts of a connection are sent on a label switching path (LSP) from an ingress node to an egress node.

We first study the feature called "streamline effect". The streamline effect is that, due to the bufferless nature of the core nodes, if some connections share a link, there will be no contention among these connections on the outgoing links at the downstream nodes. This thesis analyzes this effect and presents a loss estimation formula considering this effect. We next study the feature called "link residual capacity estimation". In IP networks, the residual bandwidth on a link is computed as the link capacity subtracted by the effective bandwidth of each connection carried. This method is not applicable to OBS networks, due to the bufferless nature. We propose a more accurate metric called residual admission capacity (RAC). We also develop a method to compute the value of RAC.

The streamline effect is used to design effective offline route optimization algorithms for best-effort traffic. We study two route optimization problems. The first problem considers the network in the normal working state where all the links are working properly. The route for each connection

is determined so as to minimize the overall network burst loss. The second problem considers the failure states apart from the normal working state. The primary and backup paths for each connection are determined in such a way to minimize the expected burst loss over the normal and failure states. The mixed linear programming (MILP) formulations and computationally efficient heuristic algorithms for the two problems are developed. The effectiveness of the algorithms is verified through numerical results obtained by solving the MILP formulations and also through simulation results on various networks.

The concept of RAC is applied to develop solutions for the problem of routing end-to-end loss guaranteed connections and two problems in configuring end-to-end loss guaranteed connections, which are the loss budget partitioning problem and the loss threshold selection problem. The loss budget partitioning problem is to choose the loss guarantee values for an end-to-end loss guaranteed connection on the links so that the end-to-end loss requirements are met and the network capacity utilization is maximized. To accomplish this, predefined loss threshold values can be associated with each link. For scalability reasons, it is desirable to have a small number of such loss thresholds. The problem of choosing such threshold values is called as loss threshold selection problem. For the routing problem, we present two algorithms, RAC based widest shortest path algorithm (RAC-WSP) and the RAC based Offline Routing algorithm (RAC-OR), for the online and offline scenarios, respectively. We also develop an RAC based loss budget partitioning (RAC-LBP) algorithm and an RAC based loss threshold selection (RAC-LTS) algorithm. The effectiveness of the proposed algorithms is verified by simulation results.

Contents

List of Figures

List of Tables

Mathmatical Notations

- ρ: the offered load.
- W: the number of wavelengths per link.
- $G(\rho, W)$: Erlang B based loss estimation formula.
- E: the expected burst loss over the normal and failure states.
- N: the number of links in the network.
- $links$: the set of links in the network. The links are number from 1 to N.
- $nodes$: the set of nodes in the network.
- $states$: the set of all the normal and failure states. The states are number from 0 to N.
- $flows$: the set of flows. Each flow is identified by a pair $< s, d >$, where s and d are the source and destination node, respectively.
- W: the number of wavelengths per link.
- $Head(s)$: the links starting from node v.
- $Tail(v)$: the links ending at node v.
- $Up(l)$: the upstream end node of link l.
- $Down(l)$: the downstream end node of link l.
- $\rho_{s,d}$: the traffic load of flow $< s, d >$.

- $x_{s,d}^k$: is 1 if the primary path of flow $< s, d >$ traverses link $k(1 \le k \le N)$, otherwise it is 0. Note that $x_{s,d}^k$ indicates the route the flow uses in *state_k*. The flow uses the backup route in *state_k* if $x_{s,d}^k = 1$ and the primary route if $x_{s,d}^k = 0$. To describe the route selection in the normal state, we additionally define $x_{s,d}^0 = 0$.

- $y_{s,d}^k$: is 1 if the backup path of flow $< s, d >$ traverses link k, otherwise it is 0.

- $a_{s,d}^{k,i}$: is 1 if the flow $< s, d >$ traverses link k in state i, otherwise it is 0.

- $\beta_{s,d}^{k,i}, \gamma_{s,d}^{k,i}$: two auxiliary boolean variables used in the definition of $a_{s,d}^{k,i}$.

- ρ_i^k: the load over link k in state i.

- $Prev(k)$: the set of the links whose downstream end node is $Up(k)$.

- $b_{s,d}^{l,k,i}$: is 1 if flow $< s, d >$ traverses the concatenation of link l and link k in state i, otherwise it is 0. Note that $l \in Prev(k)$.

- $\theta_i^{l,k}$: the load over the link concatenation of l and k in state i. Note that $n \in Prev(k)$.

- $L\_loss_i^k$: the burst loss over link k in state i.

- $Loss(state\_i)$: the burst loss in state i.

- δ: a small value (set to 10^{-8} in this chapter) that keeps the link cost greater than zero and prevents a loop in the route found.

- $\hat{G}(\rho, W)$: a piecewise linear function to approximate the non-linear $G(\rho, W)$ with interpolation.

- ρ': the traffic load of the flow whose route is to be determined.

- $\rho_E^{m,n}$: the existing traffic load over link $< m, n >$.

- $P(m)$: the node prior to node m in the shortest path from the source node to node m.

- EFL :expected loss in failure states.

- Y : the maximum node degree.

- V : the number of nodes.

- I : the total number of iterations.

- M : the umber of flows whose routes are re-computed each iteration.

- $\rho_i^{m,n}$: the traffic load of all the flows going through each link $< m, n >$ in each state i.

- $\theta_i^{P(m),m,n}$: the traffic load going through both link $< P(m), m >$ and $< m, n >$ in each state i.

- B : the number of LGTs on the link.

- $Loss(\rho, W)$: the formula to estimate the burst loss.

- $\overline{\beta}$: the residual admission capacity.

- K : the number of loss thresholds.

- H : $H = (h_1, h_2 ... \ h_K)$, the loss threshold vector.

- $(p_1, p_2 ... p_D)$: an LGT's path vector. p_m is the mth link in the path.

- γ : end-to-end loss requirement.

- σ : overall quantization cost.

- $C_{n,ij}$: the cost of customer $g_{n,j}$ being served by facility h_i in P-facility problem.

- c : average load of LGT requests.

- Z : network diameter.

- $B_{\min}^{e2e}$ and $B_{\max}^{e2e}$: the minimal and the maximal end-to-end loss guarantees provided.

- Υ_i and Ψ_i : the end-to-end loss guarantees that the i^{th} accepted LGT request required and actually provided, respectively.

- U : set of all the unadmitted LGT requests.

- A : set of all the admitted LGT requests.

- Q : an LGT request.

- R_i : minimal cost path of LGT request Q_i.

- C_i : cost of routing Q_i over path R_i.

- η : capacity stringency of an LGT request.

- φ : hop number of the shortest path.

Acronym List

ATM	asynchronous transfer mode
BORA	burst overlap reduction algorithm
DiffServ	differentiated service
EFL	expected loss in failure states
FDL	fiber delay line
FEC	forwarded equivalent class
FRR	failure recovery route
IntServ	integrated service
JET	just-enough-time
JIT	just-in-time
LAUC	latest available unscheduled channel
LAUC-VF	latest available unscheduled channel with void filling
LGT	loss guaranteed tunnel
LSP	label switching path
MILP	mixed integer linear programming
MPLS	multiple protocol label switching
NSFNET	National Science Foundation network
NSR	normal state route
OBS	optical burst switching

OCS	optical circuit switching
OPS	optical packet switching
pJET	priority just-enough-time
PPBS	probabilistic preemptive burst segmentation
QoS	quality of service
RAC	residual admission capacity
RAC-LBP	RAC based loss budget partitioning algorithm
RAC-LTS	RAC based loss threshold selection algorithm
RAC-OR	RAC based offline routing algorithm
RAC-WSP	RAC-based widest shortest path algorithm
SLFR-Heur	Streamline effect based failure recovery route optimization heuristic
SLNS-Heur	Streamline effect based normal state route optimization heuristic
SPF	shortest path first
TAG	tell-and-go
VOD	video on demand
VoIP	voice over IP
WDM	wavelength division multiplexing

Chapter 1

Introduction

Optical burst switching (OBS) [1][2][3] is an efficient switching paradigm to transmit bursty traffic over wavelength-division multiplexing (WDM) networks. It is a promising technology for the transport infrastructure of the next generation Internet. It has received a lot of research attention in the past few years.

Due to prematurity in technologies, the fiber delay lines (FDLs), which provide the buffering function in the optical domain, are still very expensive and can provide only short delays. Therefore, the core nodes in OBS networks are usually not equipped with optical buffers. It renders OBS networks new features different from the traditional IP networks. As a result, the mechanisms of routing and QoS provisioning widely used in IP networks, which are designed based on the availability of a large amount of electronic buffers at each node, are no longer efficient for OBS networks. Instead, schemes with the special features of OBS networks taken into consideration are needed. We notice due to the similarities between IP/ATM and OBS networks, usually the traditional methods can still work in OBS, but there may be better solutions with the special feature of OBS networks considered. This thesis aims to identify and analyze these special features and apply these features to design efficient schemes of connection routing and configuration for OBS networks.

1.1 Overview of OBS

WDM is a technology which effectively utilizes the huge capacity on optical fibers. With WDM, an optical fiber can carry many (tens to hundreds) non-overlapping wavelengths, each operating at the speed of a few to tens of Gbps. However, traditional WDM networks work in a circuit-switching mode where one wavelength is dedicated to one connection during the lifespan of the connection, which results in a low efficiency for the bursty data traffic. To solve this problem, optical packet switching (OPS) has been proposed, which provides better bandwidth efficiency by implementing statistical multiplexing. The processing mechanism of OPS is similar to that in the IP networks. However, OPS is not practical at present because of the technological hurdles. The main problem lies in the packet header processing which can be done only electronically instead of optically. Therefore, at every node, to remove the mismatch between the electronic processor speed and optical transmission rate, the packet payload must go through an FDL to get sufficient delay while the packet header is being processed electronically. Packet synchronization and header separation/insertion are the main hurdles.

OBS is a promising switching transmission paradigm for WDM networks. Compared with OPS, OBS is also efficient yet technologically feasible, and thus more practical. OBS networks use statistical multiplexing like OPS networks to enhance the bandwidth usage efficiency. In OBS networks, at the ingress node, data packets are assembled into large data bursts. Generally, the packets assembled into one burst are heading towards the same egress node and have the same requirements such as quality of service (QoS). Such burst assembling can help reduce the control overhead and thus improve efficiency. A control packet is sent before each data burst on a dedicated control channel along the route to the destination and is processed electronically at the core nodes to reserve an output wavelength for a period required by the data burst. As a result, data bursts can cut through the network without optical-electrical-optical (O-E-O) conversion or FDLs at the intermediate nodes. The time gap between the control packet and the data burst is set to allow for enough time for core nodes to process the control packet electronically and reserve an output wavelength for the data burst before its arrival. If a free output channel cannot be found, the data burst is dropped.

The use of bursts, instead of IP packets, as the data unit switched over the networks in OBS networks greatly reduces the amount of control overhead and the burden on the electronic devices. The separation of control packet and data burst in transmission avoids the use of expensive and large optical buffers at core nodes. Thus, OBS exploits the huge capacity of WDM networks in the optical domain and sophisticated processing capability in the electronic domain in a cost-effective way. Therefore, OBS is considered as a technology of choice for the transport infrastructure for the next generation Internet.

An OBS network is composed of core nodes, edge nodes and the WDM links. An edge node is composed of an electronic router and a burst assembler. It provides legacy interfaces and carries out the burst assembly/disassembly functions. A core node consists of optical switching matrix, switch control unit and routing and signaling processors. It is in charge of control packet processing and burst forwarding. A detailed design of these nodes was proposed by Xiong et al. [3].

Compared with the traditional IP networks, OBS networks have the following unique characteristics:

1. Bufferless or limited-buffer core nodes. Due to the high costs of FDLs, core nodes usually are not equipped with FDLs. Even if FDLs are equipped, the optical buffer can only provide very short delays up to tens of milliseconds.

2. Low delay and possibly high loss. Since there is no buffer or only limited buffer at the core nodes, a burst is simply dropped if there is no free output channel to fit it in. So, the queuing delay at the core nodes is either equal to zero (no buffer) or very small (with FDLs). As a result, the delay is not so much a concern in OBS networks. It has been shown in [4] that, even in a service differentiation scheme where the high-priority bursts get extra delay in the ingress nodes, the end-to-end delay can still meet the requirements of the most stringent real-time services. Instead, minimizing the burst loss and providing loss guarantees are much more important problems.

The traffic in OBS networks can be divided into two categories in terms of their QoS requirements. The first category is the best-effort traffic which are more tolerant to the burst loss and have no specific demand on the loss rate, while the other one is the loss-guaranteed traffic which demand the end-to-end loss rate no larger than a specific value. The first category corresponds to the non-real-time applications in the present Internet, such as web surfing and E-mail. On the other hand, the traffic of the real-time and mission critical applications, such as Voice over IP (VoIP), video on demand (VOD), live video broadcasting and video conferencing, fall into the second category.

1.2 Motivation

We address the problems of connection routing and configuration for OBS networks in this thesis. OBS networks are different from traditional IP networks in that they are usually bufferless. As a result, the traditional solutions designed for IP networks are no longer efficient and new solutions are needed. The objectives of this thesis are in two folds. First, we identify the unique features of OBS networks which arise from the bufferless nature of the core nodes. Second, we use these features to develop effective solutions for connection routing and configurations to enhance the performance of OBS networks.

We assume that the OBS networks have Multiple Protocol Label Switching (MPLS) control. We also assume that each node has full wavelength conversion, which is widely adopted in the OBS research community. We assume that the Latest Available Unscheduled Channel with Void Filling (LAUC-VF) scheduling algorithm is used, and the core nodes are not equipped with optical buffer, i.e. there is no FDL at the core nodes. Besides, no burst fragmentation or deflection routing is implemented.

Two scenarios, offline and online, are considered in this thesis. For the online scenario, we assume that the connection requests come one by one and no information of future requests is known. In the offline scenario, we assume that the traffic demand is known. The measurements in Internet traffic indicate that the aggregated load on links is quasi-stationary, which means that the network traffic statistics change relatively slowly [5]. Since bursts in OBS networks are assembled from IP streams, we expect that the traffic exhibits similar behaviors and it makes our assumption reasonable. The traffic demand may be updated from time to time, however, it is assumed that the time between two successive updates is long enough so that the traffic can be regarded as static within this period.

The term of 'burst flow' (or simply 'flow') is used to refer to the stream of bursts sent on a label switching path (LSP) from an ingress node to an egress node. Two traffic types, best-effort and loss-guaranteed, are considered in this thesis. Loss-guaranteed traffic are assumed to be carried by loss guaranteed tunnels (LGTs). An LGT is a burst flow, i.e. an LSP, with an associated end-to-end loss guarantee. LGTs are usually long-lived. They are designed and created by service providers based on the estimated traffic demand. The dynamic IP flows sent towards the same egress node with a specific loss requirement can be mapped to an appropriate LGT at the ingress node provided the total offered load is no larger than the maximum permissible load (which will be simply mentioned as 'load' in the rest of the thesis) of the LGT. An LGT is identified by its source-destination node pair and the end-to-end loss requirements. There may be more than one LGTs created between an ingress-egress node pair with different end-to-end loss requirements. In each traversing link, the loss rate of the LGT is guaranteed to be lower than a certain threshold so that the end-to-end loss requirements are met.

1.3 Contribution

In this thesis, we study two features in OBS networks which are caused by the bufferless nature of core nodes, streamline effect and residual admission capacity. We also develop effective solutions for connection routing and configurations utilizing these features.

1.3.1 Streamline Effect and its Application in Offline Route Optimization for Best-Effort Traffic

The first feature discussed in this thesis which arises from the bufferless property is the streamline effect. Traditionally Erlang B formula is used to estimate the loss over a link in OBS networks. However, the traditional method is not accurate due to ignorance of streamline effect. The streamline effect in OBS networks is that, due to the bufferless nature of the core nodes, if some flows share a link, there will be no contention among these flows on the outgoing links at the downstream nodes. This thesis analyzes this effect and presents a loss estimation formula considering the streamline effect.

We use the streamline effect to solve two offline route optimization problems for best-effort traffic. The first problem considers the case of normal working state where all the links are working properly, and a route is determined for each flow to minimize the overall burst loss. The second problem considers the failures, and the primary and backup paths for each flow are determined in such a way to minimize the expected burst loss over the normal and the failure states. We refer the first problem as the normal state route (NSR) optimization problem and the second problem as the failure recovery route (FRR) optimization problem.

For the FRR problem, we consider a failure recovery mechanism as below. For each flow, two link-disjoint LSPs, the primary LSP and backup LSP, are set up. When the network is in the normal working state, the bursts are transmitted in the primary LSP. When a link failure occurs, the end nodes of the failed link detect the failure and notify the end nodes of the failed LSPs. After receiving the notification, the source node transfers the affected flows to the pre-configured backup LSP. We assume single link failure, which has been commonly used in the literature. So when failure occurs the affected traffic can be transferred to the backup path without searching for a new route. Such a recovery scheme is fast since it is exempted from searching and setting up of a new route after a failure occurs, and it is also efficient as the routes have been optimized.

There are earlier works [6] for offline route optimization in OBS networks where the Erlang B formula is used to estimate the loss. Our work achieves better performance because we take the special feature of streamline effect into consideration. This thesis presents mixed integer linear programming (MILP) formulations for the NSR and the FRR problems. Since the MILP-based solutions are computationally intensive, heuristic algorithms are developed. The effectiveness of the algorithms is verified through numerical results obtained by solving the MILP formulations with CPLEX and also through simulation results.

1.3.2 Residual Admission Capacity and its Application in Routing and Configuring Loss Guaranteed Tunnels

Link capacity measurement is critical to routing and configuration of connections with QoS requirements. The second feature investigated in this thesis is the link residual capacity measurement. In IP networks, the residual bandwidth is computed as the link capacity subtracted by the effective bandwidth of each connection carried. The effective bandwidth of a connection is the minimal bandwidth needed to support the connection's QoS requirements. However, due to the bufferless nature of the core nodes in OBS networks, the total resource required by the aggregated connections is no longer the summation of the amount of resources required by each connection. If we adopt the residual bandwidth computation methods used for IP networks to OBS networks directly, we will obtain inaccurate results. As we will show later, the computation may show that the resource on a link is used up when there is still some capacity available to admit new connections. Also, it may show that a link has a larger residual capacity than another link where actually there is less resource available. It motivates us to develop a new method to measure the amount of residual capacity on a link in OBS networks. We propose a new metric, called *residual admission capacity (RAC)*, to measure the link residual capacity more accurately. We also develop a method to compute the value of RAC.

The concept of RAC will be used to develop solutions for the following problems in configuring LGTs:

1. Loss budget partitioning problem. It is required to choose the loss guarantee values for an

LGT on the links so that the end-to-end loss requirements are met and the network capacity utilization is maximized. It is an online per-LGT problem. This problem was considered in [7][8][9][10]. However, the algorithms proposed are either inefficient [7][8], requiring the loss guarantee values on the traversing links to be the same disregarding the different amount of residual capacity, or inflexible [9][10], demanding the loss thresholds to be uniformly distributed on a logarithmic scale. A more detailed review on these algorithms will be given in the next chapter. This thesis proposes a new algorithm, RAC-LBP (RAC based loss budget partitioning algorithm), which effectively utilizes the network capacity and it does not require the loss thresholds to follow a specific distribution. The effectiveness of RAC-LBP is verified through numerical results.

2. Loss threshold selection problem. To accomplish loss budget partitioning for an LGT, predefined loss threshold values can be associated with each link. For scalability reasons, it is desirable to have a small number of such loss thresholds. The problem of choosing such threshold values is called as loss threshold selection problem. We note that the loss threshold selection problem is an offline problem regardless of the number of LGTs and their loss requirements. On the other hand, loss budget partitioning is done for each LGT and thus is an online problem. In [9][10], the loss threshold values are assumed to be uniformly distributed on a logarithmic scale and a loss budget partitioning scheme was designed for such a loss threshold setting. This thesis proposes a new algorithm, RAC-LTS (RAC based loss threshold selection algorithm), which is more effective than the existing loss threshold selection schemes. Experiment results verify the effectiveness of RAC-LTS.

Besides, the concept of RAC will also be used to design algorithms to route LGTs. Two scenarios of routing, online and offline, are considered in this thesis. To the best of our knowledge, so far there have been no works on route selection for OBS connections with end-to-end loss guarantees. In research work about providing loss guarantee in OBS networks, usually shortest paths are assumed. The shortest path routing is very likely to create bottleneck links and reduce the network capacity utilization. This thesis presents two algorithms to route LGTs in the online and offline scenarios, respectively. For the online scenario, we develop a routing algorithm called RAC-WSP (RAC based widest shortest path algorithm). RAC-WSP is a widest shortest path (WSP) algorithm with RAC as the measurement of the residual capacity on a link. For the offline scenario, the RAC based Offline Routing algorithm (RAC-OR) is developed. Experimental results show that algorithms presented in this thesis can admit more LGT requests than other routing algorithms in the same scenario.

1.4 Organization of the Thesis

The rest of the thesis is organized as follows.

Chapter 2 reviews the research background and work related to this thesis.

Chapter 3 analyses the streamline effect and shows that the traditional Erlang B loss estimation formula is inaccurate due to this effect. A new loss estimation formula considering the streamline effect is presented. The effectiveness of the new loss estimation formula is verified by simulation results.

Chapter 4 applies streamline effect to offline route optimization for best-effort traffic in OBS networks. First we show how the consideration of streamline effect can help find a better route layout. Then the MILP formulations for the NSR and the FRR problems based on the new formula are given. As it usually costs a lot of computational resources to solve MILP formulations, we present heuristic algorithms for the two problems. We verify the effectiveness of our algorithms through numerical results obtained by solving the MILP formulations with CPLEX and also through simulation results on various networks.

Chapter 5 investigates the unique feature of OBS networks in link residual capacity measurement. We first show that inaccurate results will be obtained if we apply the traditional method of computing residual bandwidth in IP networks to OBS networks directly. Then a new metric, residual admission capacity (RAC), which measure the link residual capacity more accurately, is presented. A method to compute the value of RAC is also presented.

Chapter 6 applies the concept of RAC to design algorithms of loss budget partitioning and loss threshold selection for LGTs. First, we develop an RAC based loss budget partitioning (RAC-LBP) algorithm. We also develop an RAC based loss threshold selection (RAC-LTS) algorithm. The numerical results verify the effectiveness of RAC-LBP and RAC-LTS.

Chapter 7 presents algorithms to route LGTs. For the online scenario, RAC-WSP, which is a widest shortest path (WSP) algorithm with RAC as the measurement of the residual capacity on a link, is presented. For the offline scenario, we develop RAC based Offline Routing algorithm (RAC-OR). Experimental results show that our algorithms can admit more LGT requests than the other routing algorithms in the same scenario.

Chapter 8 concludes this thesis and suggests some directions for future work.

The works referred in this thesis are listed in **Bibliography.**

The publications based on our research are listed in **Publications.**

Chapter 2

Background and Related Work

Optical burst switching has received considerable attention in the past few years. This chapter gives a brief review on the research works on optical burst switching. It examines various aspects of OBS networks to give the background information which is relevant to the research work in this thesis.

2.1 Background of OBS

Internet has undergone an explosive growth in the past two decades. Various kinds of applications, from the non-real-time applications, such as web surfing, file transfer and E-mail, to the real-time and mission-critical applications, such as telephony, video on demand (VOD) and video conferencing, are now designed to be transmitted over the Internet. The bandwidth demand on the next generation Internet has surged in an unprecedented way. As WDM networks provide enormous transmission capacity, it will be an ideal technology of choice for the backbone of next generation Internet. In WDM networks, an optical fiber can carry many (tens to hundreds) non-overlapping wavelengths, each operating at the speed of a few to tens of Gbps. To transmit the traffic carried by Internet Protocol (IP) over WDM networks, a straight forward approach is to use a multi-layered architecture of IP-over-ATM-over-SONET-over-WDM. However, the architecture of IP-over-WDM has received much attention because it is exempted from the overheads associated with the ATM and SONET and thus the system complexity and cost are reduced.

There are mainly three optical switching techniques that have been proposed in the literature to transport IP traffic over WDM optical networks, namely OCS, OPS and OBS. OBS, as briefly reviewed in Chapter 1, combines the advantages of OCS and OPS to overcome their shortcomings to realize an all-optical switching scheme with high bandwidth utilization, high data rate, data transparency and simultaneously low complexity and cost. Therefore, OBS is a flexible and feasible solution towards the next generation optical Internet.

In OBS networks, at the ingress node, data packets are assembled into large data bursts. Generally, the packets assembled into one burst are heading towards the same egress node. A control packet is sent before each data burst on a dedicated control channel along the route to the destination and is processed electronically at the core nodes to reserve an output wavelength for a period required by the data burst. So data bursts will cut through the network without optical-electrical-optical (O-E-O) conversion or FDLs at the intermediate nodes. The time gap between the control packet and the data burst is set to allow for enough time for core nodes to process the control packet electronically and reserve an output wavelength for the data burst before its arrival. The data burst is dropped if a free output channel is not available.

As a burst is processed as a whole in each intermediate node, all the IP packets inside a burst will be treated in the same way across the network. Therefore, the performances evaluated in the packet level, such as the loss rate, should be roughly the same as that evaluated in the burst level, or per-flow level.

OBS is a link layer technique to transmit IP packets over long-haul backbone networks. It is an IP-over-WDM technology. No amendments to existing IP or TCP layer are required. The protocols of transport layer and application layer usually have the capability to deal with the possible packet loss in IP layer. Such capabilities will exert its function in an OBS based IP networks in the same manner.

As the core nodes in OBS networks usually are bufferless or limited-buffered, the queueing delay in the core nodes are equal to zero or very short, but the burst loss can be very high. Therefore, the problem of minimizing the burst loss in OBS is an important issue and has been widely studied. Another major challenge in using OBS networks as the transport infrastructure of the next generation Internet backbone is to provide support for QoS differentiation. Mission-critical and real-time applications have more stringent QoS requirements in burst loss. Much research has been done on supporting QoS differentiation in the Internet with QoS framework such as Integrated Service (IntServ)[12] and Differentiated Services (DiffServ)[13]. However, QoS mechanisms in the IP networks such as active queue management and packet scheduling are designed based on the availability of electronic buffers at the cord nodes. Therefore, many new schemes that take into consideration the unique properties of OBS networks to provide QoS are presented.

2.2 Switching Techniques of OBS

Several variants of switching techniques in OBS networks are presented in the literature, including tell-and-go(TAG) [14][15], just-in-time (JIT)[16][17] and just-enough-time (JET) [18]. These protocols differ in the way how bandwidth is reserved/released and the choice of offset time. A short description of these techniques are given below:

- In TAG protocol, the control packet is first sent on a separate channel to reserve wavelength along the path for the data burst. The data burst is transmitted on the data channel after some offset time. A control signal will be sent to release the wavelength. No acknowledgement is required for the release.
- In JIT protocol, the data burst is also transmitted after some offset time, but the wavelength is reserved immediately upon the control packet is processed. Since the control packet has no idea on the burst length, an explicit message is sent to release the bandwidth or a time-out occurs.
- In JET protocol, the control packet is sent before the data burst. The control packet contains the information of the offset time and burst length. In each intermediate node, the control packet reserves the bandwidth for the exact period that the burst will cut through.

Of the switching protocols described above, JET shows the best resource utilization efficiency as the wavelength is reserved only for the period that the burst is transmitted. Therefore, most research works in OBS nowadays assumes this protocol. This thesis takes this assumption, too. It should be noted that JET protocol is a one-way signaling protocol. Such a design reduces the transmission delay experienced by end users and the management complexity, and make OBS a technology suitable for long-haul backbone networks. However, if a control packet fails to reserve a wavelength due to contention, the burst will simply be dropped and transmission work done on the previous links is wasted. As a result, additional techniques are needed to improve the performance of JET-based OBS networks in terms of burst loss.

In order to minimize the burst loss, a different architecture, namely the time-slotted (or time-sliced) OBS networks, has been proposed and studied in the literature [19][20][21][22]. In such networks, the time in each wavelength is divided into many slots of same lengths. The IP packets are assembled into data bursts of fixed size which is equal to the length of a slot. Each burst is sent out only at the moment that a slot starts. It was shown that such an architecture can considerably reduce the burst loss [23][24]. However, it requires additional control and optical buffering at each node to synchronize every burst to align with the time slot in every traversing links.

2.3 Using MPLS for OBS

As we studied earlier in Chapter 1, we assume that OBS networks have MPLS control. Here we briefly discuss how MPLS is employed in OBS networks to realize explicit routing. In traditionally

IP networks, the packets are transferred on a hop-by-hop basis by looking up the routing table at each node using its destination address as the index. MPLS uses a different mechanism, where labels are used to make forwarding decisions at the network nodes. An OBS network with MPLS control works as follows. When an unlabeled burst enters the ingress node, the ingress node first determines the forwarding equivalence class (FEC) the burst should be in, and then puts the information of the corresponding label in the control packet. Bursts destined for a given egress node and with the same service requirement may belong to the same FEC. At an intermediate node, the label is used to determine the next hop for the control packet. The incoming label is replaced with the outgoing label which identifies the respective FEC for the downstream node. The data burst will be sent out along the path the control packet has travelled with a time gap equal to the offset time after the transmission of the control packet.

The label based forwarding method reduces the processing overhead involved in routing at the intermediate nodes, and more importantly, facilitates explicit routing and QoS control. With MPLS, explicit routing is achieved at the ingress node to set up the label switched path (LSP). Once the LSP is set up, the bursts belonging to the corresponding FEC will be forwarded along a specific route by using a label. The use of labels also allows the intermediate nodes to take some actions for meeting QoS requirement. For example, a node can control the loss rate of a flow in a link under a specific threshold by identifying the bursts belonging to a flow through its label. Such features gives OBS networks with MPLS control to provide better support for traffic engineering and QoS provisioning. Therefore, we assume that OBS networks use MPLS control in this thesis.

2.4 Techniques for Reducing Burst Loss

Since the core nodes in OBS networks do not have buffers, the delay is not a major concern. However, a burst is dropped in an OBS network if the control packet cannot find a free outgoing wavelength. With the basic paradigm of JET protocol, the burst loss rate could be very large. Therefore, several techniques, including scheduling algorithms and traffic engineering methods, have been proposed in the literature to reduce the burst loss.

2.4.1 Scheduling Algorithms

A scheduling algorithm at a node assigns a free wavelength on the outgoing link to a burst for the duration of its transmission. If FDLs are available, assignment of FDLs to a data burst is required when it cannot be immediately scheduled upon the arrival of a burst. If the scheduling algorithm is unable to find a suitable wavelength, the burst will be dropped. Therefore, a scheduling algorithm plays a vital role in burst dropping performance. Time on a wavelength is fragmented into periods occupied by bursts and periods which are idle. The idle periods are referred as voids. Scheduling algorithms always try to minimize the voids so as to pack the data bursts more tightly to make more room for the future bursts. A brief discussion on the scheduling algorithms proposed in the literature is given below.

The algorithm of Latest Available Unscheduled Channel (LAUC) [3] keeps track of the latest time any reservation will end on each wavelength, and the scheduler assigns a new data burst to the channel that will result in a minimal starting void. This algorithm is simple and has good performance in terms of the running time. However, its bandwidth efficiency is low as the void intervals are not used.

To improve the bandwidth usage efficiency, an algorithm called Latest Available Unscheduled Channel with Void Filling (LAUC-VF) [3] was proposed. LAUC-VF is similar to LAUC except that it keeps track of all void intervals (including the interval between the end time of the last burst and $+\infty$). This scheduling algorithm tries to minimize the starting void by assigning an incoming data burst to the wavelength which has a void to hold the burst and the starting void is minimal. LAUC-VF reduces the burst losses compared to LAUC but the processing time is longer.

To reduce the processing time, in [25], a set of burst scheduling algorithms utilizing the techniques from computational geometry to reduce the algorithm complexity were presented. It is shown that the average running time of this algorithm is significantly reduced while achieving the burst loss performance close to LAUC-VF.

The technique of burst rescheduling [26] was proposed to accomplish burst loss reduction by re-assigning some scheduled bursts to other wavelengths so as to make room for the incoming bursts.

Partial-drop scheduling algorithms were presented in [27][28][29][30]. Such schemes do not drop the whole burst when contention occurs. Instead, they drop only the overlapped part. As a result, the dropping rate of IP packets encapsulated in the bursts are reduced.

J. Li et al presented a set of proactive scheduling algorithms which are collectively called Burst Overlap Reduction Algorithm (BORA)[31]. The main idea of BORA is to utilize the electronic buffer at the edge nodes to delay some bursts so that the outgoing bursts are sequentialized, as a result the overlapping degree, which means the number of simultaneously arriving bursts on a node, is reduced. Accordingly the probability that the number of wavelengths is less than the number of simultaneously arriving bursts is reduced. Under this scheme, in the ingress node, a burst which overlaps with another burst will be delayed (the delay is bounded) even if there is a free wavelength to transmit it without delay.

The technique of ordered scheduling [32][33][34] was examined by some researchers. The key idea of this technique is to defer making the scheduling decision until just before the burst arrival in order to have more knowledge about other bursts. As the ordering decision is made with more information about the arrival time of other contending bursts, the bursts can be packed more tightly and a lower burst loss probability is achieved.

J. Li et al [35] identified the factors that affects the performance of a set of scheduling algorithms. It shows that the performance of any best-effort online algorithm is closely related to the range of offset time, burst length ratio, scheduling algorithm, and number of data channels. It discovers that the worst-case performance of any best-effort online scheduling algorithm is primarily determined by the maximum to minimum burst length ratio, followed by the range of offset time. Furthermore, if all bursts have the same burst length and offset time, all best-effort online scheduling algorithms generate the same optimal solution.

2.4.2 Connection Routing

Some researchers aim to improve the performance of OBS networks by configuring the connections (LSPs). Offline route optimization in OBS networks was investigated in [6]. The routes of flows are determined in such a way to minimize the overall burst loss. It is assumed that the network topology and the estimated traffic demand are known. This work uses Erlang B formula to estimate the burst loss and presents an MILP formulation and a heuristic algorithm. The problem of offline route optimization is also studied in this thesis. We argue that, with the unique feature of streamline effect taken into consideration, the performance will be better.

The technique of traffic splitting was investigated in[37] [38] [39]. Traffic splitting permits one flow to take more than one routes to the destination for better load balancing such that the overall burst loss rate is reduced. As transmitting one flow over more than one routes may cause the bursts out of order, measures must be taken to solve this problem.

2.4.3 Other Burst Loss Reduction Techniques

There are some other loss-reduction techniques that do not fall into the above two categories.

Route deflection [40] [41] is a technique where the bursts that could not find a free wavelength will not be dropped but try to find a deflected route. As the deflected route may be longer than the original route, it is possible that the data burst may arrive at a downstream node earlier than its control packet. In such cases, the data burst has to be dropped even if there are free wavelengths.

A larger offset time can mitigate this problem but will introduce an undesirable end-to-end delay. So, limited deflection routing protocol [42] is presented to reach a trade-off between burst loss reduction and limiting the end-to-end delay within an acceptable range.

M. Jin et al [43] proposed a heuristic to reduce the loss by collecting the feedback and adaptively adjusting the assembly intervals at the source nodes to reduce the burst loss.

In [44], methods combining burst fragmentation, load balancing and time slotting was presented to minimize the burst loss and achieve fairness. A scheduling algorithm combining time-slotting and fragmentation was proposed. Besides, an offset-based algorithm to provide inter-class service differentiation and intra-class fairness for the bursts belonging to classes with different levels of priority was developed.

2.5 QoS Provisioning in OBS Networks

In general, the term of QoS refers to the network's capability to provide a guaranteed and better service to selected high-priority traffic. In the literature, there are extensive research work carried out on providing QoS in IP networks. The methods of queue management and packet scheduling which are effective in IP networks cannot be directly extended to OBS networks because they are designed for the situation that each node has a large amount of electronic buffer. Therefore various QoS provisioning schemes designed for OBS networks have been developed.

The QoS schemes proposed can be divided into two groups, relative QoS and absolute QoS. The relative QoS model provides relative service quality ordering among classes. There is no service level guarantee (which is usually the edge-to-edge loss guarantee in the case of OBS) in each class. On the other hand, the absolute QoS model guarantees the service level of users, namely the edge-to-edge loss rate is guaranteed to be no larger than a specific value.

2.5.1 Relative QoS

The relative QoS problems proposed in the literature can be further classified into two categories. The first category is the qualitative service differentiation, where there is no quantitative target to reach. The only assurance from the network operator is that a higher class will receive a better service than a lower class. The other category is the model of proportional QoS which provides quantitative service differentiation. For two classes, i and j, A proportional QoS scheme aims to control the burst loss rate of the two classes, CL_i and CL_j, following the relationship that $CL_i/CL_j = \kappa_i/\kappa_j$, where κ_i and κ_j are predefined parameters. The model of proportional QoS has drawn more attention from the research community as it provides the network operator the ability to adjust the performance spacing between different classes quantitatively.

2.5.1.1 Qualitative Service Differentiation

The extra-offset-time-based pJET scheme was proposed in [4] [46]. In this scheme, a high-priority data burst will get extra delay at the source node to have a larger offset time. This scheme is based on the observation that a control packet is more likely to find free wavelengths if the offset time is larger. This scheme is simple as it requires no extra control in the core nodes. However, the high priority bursts will have a larger end-to-end delay.

The scheme of differentiated scheduling [45] was designed to support differentiated services by dynamically choosing the extra offset time. In this scheme, each OBS node can adjust the data burst loss rates for different classes of bursts and satisfy differentiated QoS requirement with the available resources.

Preemption based service differentiation schemes were proposed in [47][48][49]. In [47][48], strict service differentiation is achieved by allowing high priority bursts to preempt low priority burst. A probabilistic preemptive schemes was proposed in [49]. In this scheme, a high-priority burst will preempt a low priority burst in a probabilistic manner when no free wavelength is available. The probability of preemption can be adjusted to get different service differentiation ratios.

A burst segmentation based service differentiation scheme was proposed in [30]. It preferentially segments and deflects bursts of different priorities when contentions occurs. A high priority burst is allowed to preempt a segment of the low-priority bursts. The preempted part of the low-priority burst is not dropped, but deflected. Alternatively, packets of different priorities can be can be assembled into a burst in a decreasing priority from head to tail. When contention occurs, the low priority segment of the contending burst is preempted. Under the schemes proposed in this work, after the segmentation, a new control packet has to be generated with the updated burst length information and sent downstream to release the unnecessary wavelength duration reserved for the burst before segmentation. It incurs extra signalling overhead.

2.5.1.2 Proportional QoS

Proportional drop policy was proposed in [50] to achieve proportional QoS between traffic classes. This mechanism keeps track of the loss rate of each class in each link and choose to drop some bursts intentionally to meet the requirements of proportional QoS on each link. This scheme may incur excessive droppings and results in a low bandwidth efficiency. A preemptive scheduling technique was presented in [51] to realize the proportional QoS on a link. This scheme allows the high-priority bursts to partially preempt the low-priority bursts.

A problem with the above two schemes is that they aim to provide per-hop proportional QoS only. However, per-hop proportional QoS does not lead to end-to-end QoS proportionally. A scheme to select the offset time of each class adaptively based on the feedback to achieve end-to-end proportional QoS was proposed in [52]. In [53], it was proposed to combine wavelength preemption and bandwidth allocation to schedule bursts at the core nodes to provided end-to-end proportional QoS. A Probabilistic Preemptive Burst Segmentation (PPBS) scheme was proposed in [54] to enable high priority bursts to preempt and segment low priority bursts in a probabilistic fashion. It can achieves 100% isolation among priority classes, and burst loss probabilities can be controlled by tunable parameters. The author also shows that PPBS scheme can achieve proportional loss differentiation.

2.5.2 Absolute QoS

For many real-time and mission critical services, such as video conferencing, voice over IP (VoIP) etc., the damage caused by burst losses cannot be mitigated by retransmission. In [55], Maach A. et al considered burst retransmission to mitigate the impact of burst loss on end users. However, the author shows that such a scheme is suitable for networks of smaller size, such as the metropolitan networks only, not for long-haul backbone network. We note that OBS is considered as a promising technology for the next generation backbone networks due to its extremely large capacity. Therefore, absolute QoS service which provides bounded end-to-end loss rate is critical to multi-service OBS networks. The following gives a reviews absolute QoS schemes in OBS networks.

To provide end-to-end loss guarantee to an OBS connection, the following components are indispensable:

1. a mechanism to guarantee the loss rate of the connection no larger than a certain value on each traversing link.
2. a method of loss budget partitioning. This method determines or chooses the value of guaranteed loss rate on each traversing link so that the accumulative guaranteed loss rates along the path is no larger than the end-to-end loss guarantee required by the users.

Besides, due to the scalability problem, the possible loss guarantee values on the links can only take a finite number of values. As a result, it is quite likely that the end-to-end loss guarantee granted is not equal to, but smaller than, the required value. Loss thresholds are the possible loss guarantee values supported on a link. Improper loss threshold setting will cause unnecessary capacity consumption and reduces the network utilization efficiency. Loss threshold selection problem studies how to select a set of loss thresholds so that the network capacity is utilized efficiently.

So, to provide end-to-end loss guarantees in OBS network, the problems of providing loss guarantee on a link, loss budget partitioning and loss threshold selection need to be solved. A brief review on the schemes proposed to solve the three problem in the literature is given below.

2.5.2.1 Providing Loss Guarantee on a Link

The schemes proposed in the literature to guarantee the loss rate of an OBS connection lower than a specific threshold can be divided into three categories. They all classify the loss guaranteed connections on a link into a number of groups, with each group having a loss threshold. The loss rate a connection undergoes on a link is guaranteed to be no larger than the loss threshold of the group it belongs to. These schemes work as follows:

1. Wavelength grouping scheme [7][11]: this scheme divides the wavelengths on a link into several sets and allocate one set to a group. Incoming bursts of a particular class can only be scheduled within the wavelengths allocated for this group. A dynamic wavelength grouping scheme was proposed in [7]. In this scheme, the node keeps track of the loss rate of the groups. If the predefined threshold of a specific group is violated twice consecutively, the node will expand the corresponding wavelength set by moving a wavelength from a group which does not have two consecutive threshold violations. A sharing grouping scheme was proposed in

[11], where some wavelengths can be shared by different groups to enhance the statistical multiplexing.

2. Early Drop Scheme [8]: This scheme drops the low-priority bursts when the loss rate of the high-priority traffic is close to the predefined threshold.

3. Preemptive scheme [9][10]: in this scheme, when a control packet fails to reserve an output wavelength, a contention list is built, which includes the incoming burst and the bursts it overlaps. Within the contention list, the burst belonging to the connection which is farthest from breaching the loss threshold is dropped. In this way, the distance of the actual loss rate of each connection to the corresponding loss threshold is kept almost equal.

Of the above schemes, the first scheme reduces the level of statistical multiplexing, so the load that can be carried is less. The second scheme may drop low-priority bursts unnecessarily when they are not in contention with high-priority bursts. Comparatively, the preemptive scheme achieves the highest efficiency.

2.5.2.2 Loss Budget Partitioning

The loss budget partitioning problem is to choose the loss guarantee values for an LGT on the links so that the end-to-end loss requirements are met and the network capacity utilization is maximized. To accomplish this, predefined loss threshold values can be associated with each link. For scalability reasons, it is desirable to have a small number of such loss thresholds. The problem of choosing such threshold values is called as loss threshold selection problem. We note that the loss threshold selection problem is an offline problem regardless of the number of LGTs and their loss requirements. On the other hand, loss budget partitioning is done for each LGT and thus is an online problem.

For the loss budget partitioning problem, there are three schemes proposed in the literature.

1. Average loss budget partitioning was proposed in [8]. In this scheme, the loss guarantee for an OBS connection on a link is calculated by equally dividing the end-to-end loss requirements by the maximum network diameter.
2. In the scheme proposed in [9][10], a less stringent guarantee is assigned at a more resource-stringent link and a more stringent guarantee at a less stringent link. However, this scheme requires the thresholds be uniformly distributed on a logarithmic scale. The details of the algorithm can be found in [9][10].

Among the two schemes, the second scheme performs better in resource utilization. However, it requires the loss thresholds follow a specific distribution.

2.5.2.3 Loss Threshold Selection

Schemes proposed in the literature vary depending on whether loss threshold selection is needed:

1. In [7][8], the loss guarantees of a connection on the links are decided by the end-to-end loss requirements and they could not take different values on different links. So a loss threshold

selection method is not needed.

2. In [9][10], the loss guarantee values of an OBS connection could be different according to the different amount of residual capacity on the traversing links. In such schemes, different sets of loss thresholds will lead to different capacity utilization efficiency. As a result, loss threshold selection becomes necessary.

   In [9][10], it is suggested to set the thresholds uniformly distributed on a logarithmic scale. Hence, the loss thresholds are chosen as h_1, αh_1, $\alpha^2 h_1$,..., $\alpha^{K-1} h_1$, assuming that there are K thresholds.

Of the two categories above, the latter yields better resource utilization as it considers the remaining resources in assigning loss guarantees. So the study of loss threshold selection is necessary for OBS networks of higher network utilization efficiency. Therefore, we will study the problem of loss threshold selection in this thesis.

As reviewed above, there have been a lot of approaches presented in the literature for OBS networks. Future developments in optical technologies and end users' demand for bandwidth will decide which approaches are more practical to implement in the real world. Such key technologies include the optical buffers, processing speed of switching processors and the cost of optical fiber fabrication and extension, the wavelength density of WDM links, etc. Their developments will change the ratios between the cost and performance of different approaches. On the other hand, the end users' demand for high-bandwidth applications is another important factor in approach selection in practical networks. If applications consuming a large amount of bandwidth are demanded in the market, more complicated approaches, where the bandwidth is utilized more effectively, will be preferred. Otherwise simple and less costly approaches will be chosen.

2.6 Summary

This chapter presented a brief survey on research work on optical burst switching. First, we introduced the background of OBS networks and the switching techniques. Then various techniques to reduce the burst loss were reviewed. Next, we looked into the problems of QoS provisioning. We described two schemes, the relative service differentiation and absolute QoS guarantee. We reviewed solutions for these two schemes.

Chapter 3

Streamline Effect

In this chapter we analyze the streamline effect which arises from the special features of bufferless core nodes in the OBS networks. We then develop a loss estimation formula with this effect considered.

3.1 Streamline Effect and Loss Estimation

The Erlang B formula is usually used to estimate burst loss in OBS networks. The Erlang B formula assumes that all the flows are independent and contend with each other. However, since there is no buffering at the core nodes in OBS networks, if two or more flows share more than one consecutive link along their paths, the relative temporal relationship among the bursts of these flows will not change along these links. As a result, the contention among these flows can only take place at the

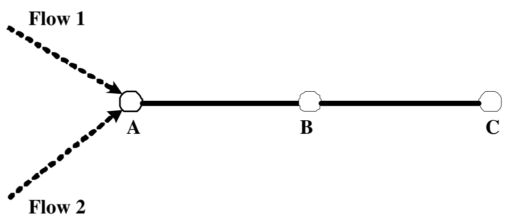

Figure 3.1: Illustration of the Streamline Effect

first shared link. Due to this effect, the loss on a subsequent link is likely to be less than that computed by the Erlang B formula. Consider the example shown in Figure 3.1. Suppose that two flows merge at node A and share link AB and link BC. There should be no loss on link BC, because link AB has removed all the contentions between the two flows. However, the estimated loss given by the Erlang B formula below is greater than zero.

$$G(a, W) = a \times ErlangB(a, W) = a(\frac{\frac{a^W}{W!}}{\sum_{m=0}^{W} \frac{a^m}{m!}}) \tag{3.1}$$

where W is the number of wavelengths per link and a is the offered load on link BC ($a = \lambda/\mu$, where λ and $1/\mu$ are the arrival rate and the mean burst length, respectively. The corresponding normalized load can be derived by dividing a by W.). Therefore, to estimate the loss more accurately, we need to take the following effect into consideration: if some flows share a link, there will be no contention among these flows if they traverse the same next downstream link. We call this effect as the *streamline effect*.

We now derive the loss estimation formula with the streamline effect taken into consideration.

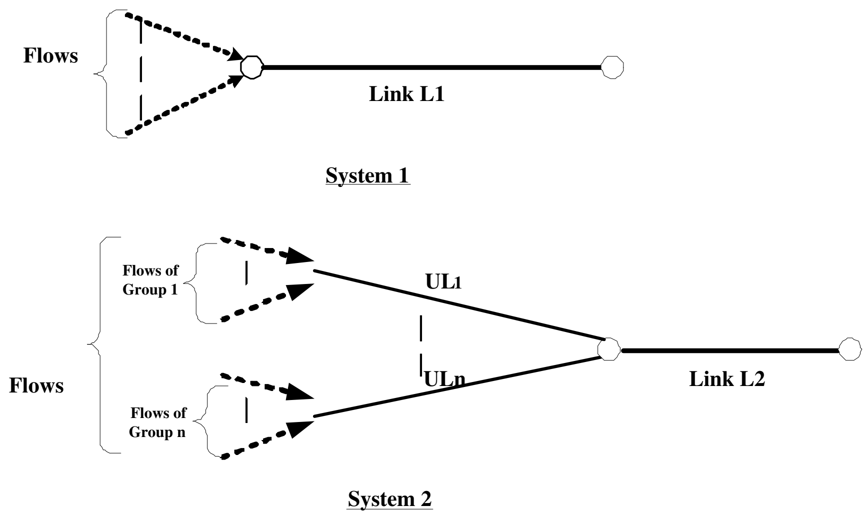

Figure 3.2: Comparison of Two Systems

We note that an intuitive explanation of the streamline effect has been given in [56], and here we give a detailed mathematical analysis. Consider two systems with the same input flows, as shown in Figure 3.2. In System 1, all the flows enter link L1 from different input links and are independent, which is the case assumed by the Erlang B formula. In System 2, flows are divided into n groups, and the flows of the ith group merge at the ith input link UL_i before reaching link L2. System 2 is the case in OBS networks. The offered load of the ith group is given by ρ_i. The total offered load is given by $\rho = \sum_{i=1}^{n} \rho_i$.

First we show that the total losses in the two systems are approximately equal when W tends to be large. The throughput in System 1, denoted by Γ_1, is

$$\Gamma_1 = \rho - G(\rho, W) = \rho\left(\frac{\sum_{m=0}^{W-1} \frac{\rho^m}{m!}}{\sum_{m=0}^{W} \frac{\rho^m}{m!}}\right) \tag{3.2}$$

In System 2, suppose that the propagation delay of the ith input link UL_i is τ_i. When a burst arrives at the jth input link at time t, the number of bursts being served at the ith input link at time $t + \tau_j - \tau_i$ is S_i. So we have the following observations:

1. If $(\sum_{i=1}^{n} S_i) < W$, the newly-arrived burst survives since it can always find a free wavelength either on the input link UL_j or link L2.
2. If $(\sum_{i=1}^{n} S_i) \geqslant W$, the newly-arrived burst is dropped. If $S_j = W$, it will be dropped on the input link UL_j. If $S_j < W$, but $(\sum_{i=1}^{n} S_i) \geqslant W$, it will be dropped on Link L2.

As a result, the throughput in System 2, denoted by Γ_2, is equal to $\rho \times \Pr((\sum_{i=1}^{n} S_i) < W)$. Because the traffic in each input link K_i are independent, and the burst arrival process is Poisson, we have

$$\begin{aligned}
\Gamma_2 &= \rho \times \Pr(\sum_{i=1}^{n} S_i < W) \qquad (3.3)\\
&= \rho \times (\sum_{m=0}^{W-1} (\sum_{S_1+S_2\ldots+S_N=m, 0\leqslant S_i \leqslant W} (\prod_{i=1}^{n} \Pr(S_i)))) \\
&= \rho \times (\sum_{m=0}^{W-1} (\frac{\sum_{S_1+S_2\ldots+S_N=m, 0\leqslant S_i \leqslant m} (\prod_{i=1}^{n} \frac{{\rho_i}^{S_i}}{S_i!})}{\prod_{i=1}^{n} (\sum_{k=0}^{W} \frac{{\rho_i}^k}{k!})}))
\end{aligned}$$

According to the formula,

$$(\sum_{i=1}^{n} a_i)^M = \sum_{b_1+b_2\ldots+b_N=M, 0\leqslant b_i \leqslant M} (\prod_{i=1}^{n} \frac{M!(a_i)^{b_i}}{b_i!}) \qquad (3.4)$$

we have

$$\sum_{S_1+S_2..+S_N=m, 0\leqslant S_i\leqslant m} (\prod_{i=1}^{n} \frac{{\rho_i}^{S_i}}{S_i!}) = \frac{(\sum_{i=1}^{n} \rho_i)^m}{m!} = \frac{\rho^m}{m!} \tag{3.5}$$

Using Eq(3.5) in Eq(3.3), we have

$$\Gamma_2 = \rho \times (\sum_{m=0}^{W-1} (\frac{\frac{\rho^m}{m!}}{\prod_{i=1}^{N} (\sum_{k=0}^{W} \frac{{\rho_i}^k}{k!})})) \tag{3.6}$$

Since $\sum_{k=0}^{\infty} \frac{{\rho_i}^k}{k!} = e^{\rho_i}$, for a large value of W (which is the usual case in OBS networks), we have $(\sum_{k=0}^{W} \frac{{\rho_i}^k}{k!}) \approxeq e^{\rho_i}$. So,

$$\Gamma_2 \approxeq \rho \times (\sum_{m=0}^{W-1} (\frac{\frac{\rho^m}{m!}}{\prod_{i=1}^{n} e^{\rho_i}})) = \rho \times (\sum_{m=0}^{W-1} (\frac{\frac{\rho^m}{m!}}{e^{\rho}})) \tag{3.7}$$

Using the same approximation technique in Eq(3.7), we have

$$\Gamma_1 \approxeq \rho \times (\sum_{m=0}^{W-1} (\frac{\frac{\rho^m}{m!}}{e^{\rho}})) \tag{3.8}$$

The above analysis shows $\Gamma_1 \approxeq \Gamma_2$. So the overall burst loss in System 2 is also $G(\rho, W)$. Since the burst loss at the ith input link K_i is $G(\rho_i, W)$, the loss at Link L2 is given by:

$$G(\rho, W) - \sum_{i=1}^{n} G(\rho_i, W) \tag{3.9}$$

Eq(3.9) gives the loss estimation formula considering the streamline effect. The loss is equal to the loss estimated by the Erlang B formula with the contentions between the flows from the same upstream link removed. We can see that the new formula gives the correct loss estimation in the case illustrated in Figure 3.1.

3.2 Numerical Results

We validate the streamline loss estimation formula on two networks. First, we validate the formula on a 6-node network shown in Figure 3.3. We consider three flows.

Flow 1: A->C->E->F

Flow 2: B->C->E->F

Flow 3: D->E->F

The loads of the Flow 1 and Flow 2 are 0.3, respectively, and the load of Flow 3 is varied from 0 to 0.15. Figure 3.4 compares the loss rates on Link EF given by the simulation and that estimated by the Erlang B formula and the streamline formula. We can observe that the estimation given by the streamline formula is closer to the simulation result than that by the Erlang B formula.

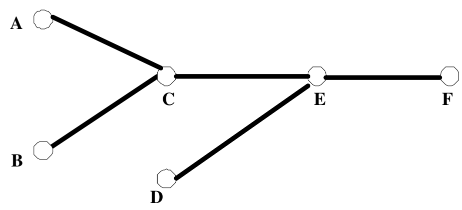

Figure 3.3: A 6-Node Netwok

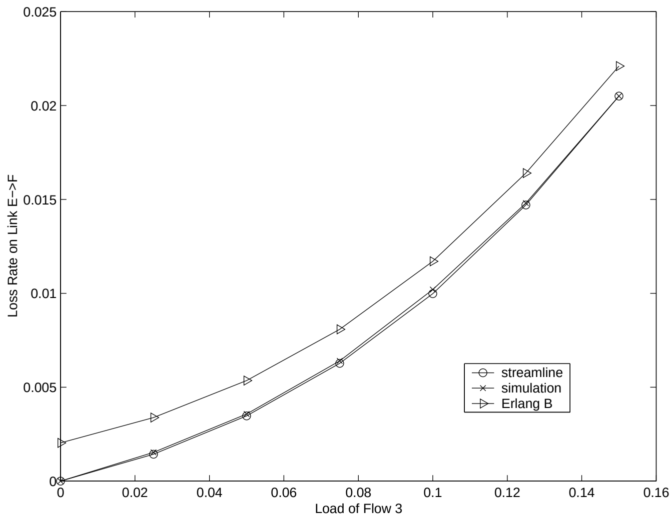

Figure 3.4: Comparions of Burst Loss Rates Estimated by Different Formulas in the 6-Node Network (Fig 3.3)

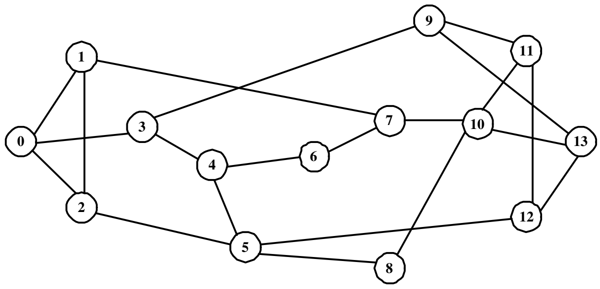

Figure 3.5: NSFNET Topology

Next, we validate the formula on the NSFNET network shown in Figure 3.5. We assume there is a flow between each node pair. We assume that each flow takes the shortest path. The load of each flow (except the one from node 11 to node 1) is randomly chosen between 0 and 0.32. We vary the load of the flow from node 11 to node 1 and observe the end-to-end loss rate of this flow. Figure 3.6 shows the result obtained from the simulation, estimated by Erlang B formula and estimated by the streamline formula. We can observe that the estimation given by the streamline formula is closer to the simulation result.

3.3 Summary

In this chapter, we have discussed the streamline effect, a special feature of OBS networks which arises from the bufferless nature of the core nodes. A new loss estimation formula which considers the streamline effect has been developed. The simulation results showed this formula is more accurate in estimating the loss rate than the Erlang B formula.

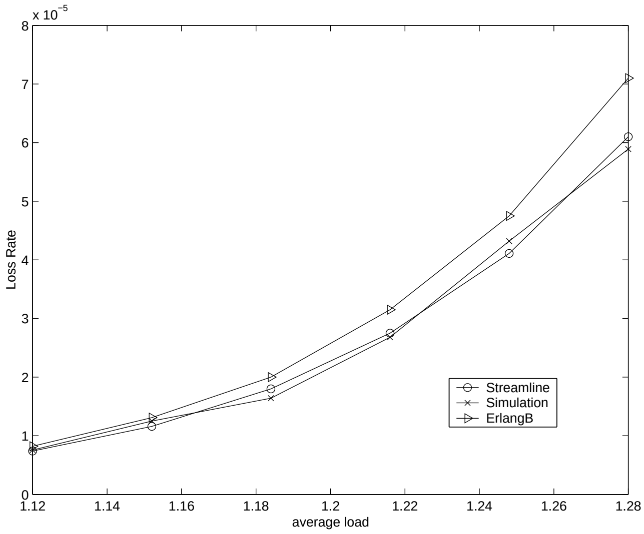

Figure 3.6: Comparions of Burst Loss Rates Estimated by Different Formulas in the NSFNET Network (Fig 3.5)

Chapter 4

Offline Route Optimization Considering Streamline Effect

This chapter considers special feature of streamline effect in OBS networks for offline route optimization of best-effort traffic. Two route optimization problems are studied. The first problem considers the usual case of normal state where all the links are working properly, and one route is determined for each flow to minimize the overall burst loss. The second problem considers the failures, and the primary and backup paths for each flow are determined in such a way to minimize the expected burst loss over the normal and the failure states. We refer the first problem as the normal state route (NSR) optimization problem and the second problem as the failure recovery route (FRR) optimization problem.

For the FRR problem, we consider a failure recovery mechanism as below. For each flow, two link-disjoint label switched paths (LSPs), the primary LSP and backup LSP, are set up. When the network is in the normal working state, the bursts are transmitted through the primary LSP. When a link failure occurs, the end nodes of the failed link detect the failure and notify the end nodes of the failed LSPs. After receiving the notification, the source node shifts the affected flows to the pre-configured backup LSP. We assume single link failure, which has been commonly used in the literature. So when a failure occurs the affected traffic could be transferred to the backup path without searching for a new route. Such a recovery scheme is fast since it is exempted from the searching and setup of a new route after a failure occurs, and it is also efficient as the routes have been optimized. There have been some works done on OBS fault managements [59][60][61][62], but they do not consider the problem of primary/backup route selection to minimize the expected burst loss.

The problem of protection route layout optimization for WDM circuit-switching networks has been well studied [63][64][65][66]. However, to the best of our knowledge, there is no such works done for OBS networks. The protection route layout schemes for WDM circuit switching networks cannot be extended readily to OBS networks, mainly due to the different switching architecture and optimization objectives. In circuit switching WDM networks, there will not be data loss on a connection once it is set up. So its optimization aims to minimize the resource consumption, namely the number of wavelengths assigned. On the contrary, OBS is a transmission paradigm using statistical multiplexing, where the bursts may be discarded during the transmission and the loss rate may be high if the routes are not properly designed. Therefore, a key objective for the route layout optimization in OBS network is to reduce the mean burst loss rate.

Offline route optimization in OBS networks has been studied in [6] wherein the Erlang B formula is used to estimate the loss. Our work differs in that we consider the special feature of streamline effect.

We will first explain how the use of streamline effect can help find a better route layout in OBS networks. Then the mixed integer linear programming (MILP) formulations for the NSR and the FRR problems are developed. Because of the intensive computation needed to solve MILP formulations, heuristic algorithms are developed.

4.1 Impact of Streamline Effect on Route Optimization

With the new loss estimation formula, a better route layout can be found by making a good use of the streamline effect. The Erlang B based route optimization minimizes the summation of losses on the links traversed estimated by the Erlang B formula. The load-balancing route optimization would make the loads distributed as balanced as possible. However, such route layouts may not be optimal in OBS networks due to the streamline effect. Consider the network topology shown in Figure 4.1. Assume that there are four flows, with the loads as shown, all destined for node F. Note that these flows have no choice on the part of routes before node C, so we focus on the part of route selection after node C.

With the Erlang B formula based or the load-balancing route optimization, the following route layout is selected. This layout makes the loads in the route C->D->F and C->E->F equal, and the summation of the Erlang B estimated loss in the links of CD, CE, DF and EF is minimized.

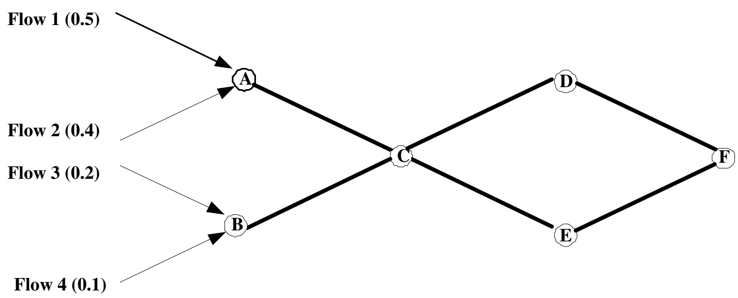

Figure 4.1: Illustration of the Benefit of Considering the Streamline Effect in Route Optimization

Flow 1: A->C->D->F

Flow 2: A->C->E->F

Flow 3: B->C->E->F

Flow 4: B->C->D->F

With the new loss estimation formula, the following layout is selected. In this layout, the summation of the loss estimated by the new formula is minimized. There is no contention between Flow 1 and Flow 2 after node C due to the streamline effect, and thus they take the same route after node C. So are Flow 3 and Flow 4.

Flow 1: A->C->D->F

Flow 2: A->C->D->F

Flow 3: B->C->E->F

Flow 4: B->C->E->F

For the two layouts, the burst losses on link AC and link BC are the same. The burst loss on the links of CD, DF, CE and EF in the first layout is larger than zero. However, in the second layout, the burst loss over these links is 0, because of the streamline effect. Therefore, the route layout obtained by the streamline effect based loss formula results in lower burst loss.

4.2 The MILP formulation

In this section, we develop MILP formulations for the NSR and the FRR problems based on the new loss estimation formula.

The NSR problem can be stated as follows:

Given an OBS network and a traffic demand, it is required to determine a route for each flow so as to minimize the overall burst loss.

The FRR problem can be stated as follows:

Given an OBS network and a traffic demand, it is required to determine a pair of link-disjoint primary and backup paths for each flow to minimize the expected burst loss over the normal and failure states. The expected burst loss, E, is defined as below:

$$E = \sum_{i=0}^{N} \Pr(state\_i) \times Loss(state\_i)$$

where N is the number of links in the network, and the links are numbered from 1 to N.

state_0 is the state that all the links are in working conditions. We call this state as the normal state. In the normal state, all the flows use the primary paths.

$state\_i (1 \leqslant i \leqslant N)$ is the state where the ith link fails. We call these states as failure states. In the failure states, the flows whose primary paths are affected by the failed link switch their traffic to the backup path, while those flows whose primary paths are not affected continue to send their traffic over their primary paths.

$\Pr(state\_i)$ is the probability that the network is in *state_i*. The probability of each state can be estimated based on historical statistics. For example, a link with more failures in the past is more prone to fail in the future and thus should be assigned a larger value of $\Pr(state\_i)$. The estimation can also be done according to the physical and geographical conditions of the links. For example, a longer link is more likely to suffer from failures. In this report, we suppose that such estimation have been done, and values of $\Pr(state\_i)$ are all known.

$Loss(state\_i)$ is the overall burst loss in *state_i*.

4.2.1 Notation

- *links*: the set of links in the network. The links are number from 1 to N.
- *nodes*: the set of nodes in the network.
- *states*: the set of all the normal and failure states. The states are number from 0 to N.

- $flows$: the set of flows. Each flow is identified by a pair $< s, d >$, where s and d are the source and destination node, respectively.

- W: the number of wavelengths per link.

- $Head(v)$: the links starting from node v.

- $Tail(v)$: the links ending at node v.

- $Up(l)$: the upstream end node of link l.

- $Down(l)$: the downstream end node of link l.

- $\rho_{s,d}$: the traffic load of flow $< s, d >$.

- $x^k_{s,d}$: is 1 if the primary path of flow $< s, d >$ traverses link $k(1 \leq k \leq N)$, otherwise it is 0. Note that $x^k_{s,d}$ indicates the route the flow uses in $state\_k$. The flow uses the backup route in $state\_k$ if $x^k_{s,d} = 1$ and the primary route if $x^k_{s,d} = 0$. To describe the route selection in the normal state, we additionally define $x^0_{s,d} = 0$.

- $y^k_{s,d}$: is 1 if the backup path of flow $< s, d >$ traverses link k, otherwise it is 0.

- $a^{k,i}_{s,d}$: is 1 if the flow $< s, d >$ traverses link k in state i, otherwise it is 0.

- $\beta^{k,i}_{s,d}, \gamma^{k,i}_{s,d}$: two auxiliary boolean variables used in the definition of $a^{k,i}_{s,d}$.

- ρ^k_i: the load over link k in state i.

- $Prev(k)$: the set of the links whose downstream end node is $Up(k)$.

- $b^{l,k,i}_{s,d}$: is 1 if flow $< s, d >$ traverses the concatenation of link l and link k in state i, otherwise it is 0. Note that $l \in Prev(k)$.

- $\theta_i^{l,k}$: the load over the link concatenation of l and k in state i. Note that $n \in Prev(k)$.

- $L\_loss_i^k$: the burst loss over link k in state i.

- $Loss(state\_i)$: the burst loss in state i.

- δ: a small value (set to 10^{-8} in this chapter) that keeps the link cost greater than zero and prevents a loop in the route found.

- $\hat{G}(\rho, W)$: a piecewise linear function to approximate the non-linear $G(\rho, W)$ with interpolation.

$$\hat{G}(\rho, W) = \frac{G(\rho_m, W) - G(\rho_{m-1}, W)}{\rho_m - \rho_{m-1}}(\rho - \rho_{m-1}) + G(\rho_{m-1}, W) \quad \rho_m > \rho \geq \rho_{m-1}, m = 1, 2, , ...K \tag{4.1}$$

4.2.2 MILP1: NSR Problem Formulation

Objective: Minimize the burst loss in the normal state $Loss(state\_0)$

Constraints: The problem is subject to the following constraints:

1. Flow conservation demand.

$$\sum_{k \in tail(v)} x^k_{s,d} - \sum_{k \in head(v)} x^k_{s,d} = \begin{cases} 1 & \text{if } v = s \\ -1 & \text{if } v = d \\ 0 & otherwise \end{cases} \qquad \forall < s, d > \in flows, \forall v \in nodes \tag{4.2}$$

2. By definition, $b^{l,k,0}_{s,d} = x^k_{s,d} \times x^l_{s,d}$. However, this expression is non-linear. So the following linear constraints are defined. They give the same results when all the variables are boolean.

$$b^{l,k,0}_{s,d} \leq (x^k_{s,d} + x^l_{s,d})/2 \qquad \forall < s, d > \in flows, \forall k \in links, \forall l \in prev(k) \tag{4.3}$$

$$b^{l,k,0}_{s,d} \geq (x^k_{s,d} + x^l_{s,d})/2 - 0.5 \qquad \forall < s, d > \in flows, \forall k \in links, \forall l \in prev(k) \tag{4.4}$$

3. The definition of $\rho^k_0, \theta^{l,k}_0$ and $L\_loss^k_0$.

$$\rho^k_0 = \sum_{<s,d> \in F} \rho_{s,d} \times x^k_{s,d} \qquad \forall k \in links \tag{4.5}$$

$$\theta^{l,k}_0 = \sum_{<s,d> \in F} \rho_{s,d} \times b^{l,k,0}_{s,d} \qquad \forall k \in links, \forall l \in prev(k) \tag{4.6}$$

$$L\_loss^k_0 = \delta + \hat{G}(\rho^k_0, W) - \sum_{l \in prev(k)} \hat{G}(\theta^{l,k}_0, W) \qquad \forall k \in links \tag{4.7}$$

$$Loss(state\_0) = \sum_{k \in links} L\_loss^k_0 \tag{4.8}$$

4.2.3 MILP2: FRR Problem Formulation

Objective: Minimize the expected burst loss E

Constraints: The problem is subject to the following constraints:

1. Flow conservation demand for the primary and the backup routes.

$$\sum_{k \in tail(v)} x^k_{s,d} - \sum_{k \in head(v)} x^k_{s,d} = \begin{cases} 1 & \text{if } v = s \\ -1 & \text{if } v = d \\ 0 & otherwise \end{cases} \quad \forall < s,d > \in flows, \forall v \in nodes \tag{4.9}$$

$$\sum_{k \in tail(v)} y^k_{s,d} - \sum_{k \in head(v)} y^k_{s,d} = \begin{cases} 1 & \text{if } v = s \\ -1 & \text{if } v = d \\ 0 & otherwise \end{cases} \quad \forall < s,d > \in flows, \forall v \in nodes \tag{4.10}$$

2. The primary and backup paths are link-disjoint.

$$x^k_{s,d} + y^k_{s,d} \leq 1 \qquad \forall < s,d > \in flows, \forall k \in links \tag{4.11}$$

3. By definition, $a^{k,i}_{s,d} = x^k_{s,d} \times (1 - x^i_{s,d}) + y^k_{s,d} \times x^i_{s,d}$. We use the following linear constraints to replace the multiplication expression. These constraints give the same results if the variables involved are all boolean.

$$\beta^{k,i}_{s,d} \leq (x^k_{s,d} + 1 - x^i_{s,d})/2 \qquad \forall < s,d > \in flows, \forall k \in links, \forall i \in states \tag{4.12}$$

$$\beta^{k,i}_{s,d} \geq (x^k_{s,d} + 1 - x^i_{s,d})/2 - 0.5 \qquad \forall < s,d > \in flows, \forall k \in links, \forall i \in states \tag{4.13}$$

$$\gamma^{k,i}_{s,d} \leq (y^k_{s,d} + x^i_{s,d})/2 \qquad \forall < s,d > \in flows, \forall k \in links, \forall i \in states \tag{4.14}$$

$$\gamma^{k,i}_{s,d} \geq (y^k_{s,d} + x^i_{s,d})/2 - 0.5 \qquad \forall < s,d > \in flows, \forall k \in links, \forall i \in states \tag{4.15}$$

$$a^{k,i}_{s,d} = \beta^{k,i}_{s,d} + \gamma^{k,i}_{s,d} \qquad \forall < s,d > \in flows, \forall k \in links, \forall i \in states. \tag{4.16}$$

4. We have $b^{l,k,i}_{s,d} = a^{k,i}_{s,d} \times a^{l,i}_{s,d}$ by definition. By the same technique as above, we use the following linear constraints to replace the non-linear multiplication expression.

$$b^{l,k,i}_{s,d} \leq (a^{k,i}_{s,d} + a^{l,i}_{s,d})/2 \qquad \forall < s,d > \in flows, \forall k \in links, \forall i \in states \tag{4.17}$$

$$b^{l,k,i}_{s,d} \geq (a^{k,i}_{s,d} + a^{l,i}_{s,d})/2 - 0.5 \qquad \forall < s,d > \in flows, \forall k \in links, \forall i \in states \tag{4.18}$$

5. The definition of ρ^k_i and $\theta^{l,k}_i$.

$$\rho^k_i = \sum_{< s,d > \in F} \rho_{s,d} \times a^{k,i}_{s,d} \qquad \forall < s,d > \in flows, \forall k \in links, \forall i \in states \tag{4.19}$$

$$\theta^{l,k}_i = \sum_{< s,d > \in F} \rho_{s,d} \times b^{l,k,i}_{s,d} \qquad \forall k \in links, \forall l \in prev(k), \forall i \in states \tag{4.20}$$

6.The definition of $L\_loss^k_i, Loss(state\_i)$ and E.

$$Loss_i^k = \delta + \hat{G}(\rho_i^k, W) - \sum_{l \in prev(k)} \hat{G}(\theta_i^{l,k}, W) \qquad \forall k \in links, \forall i \in states \tag{4.21}$$

$$Loss(state\_i) = \sum_{k \in links} Loss_i^k \qquad \forall i \in states \tag{4.22}$$

$$E = \sum_{i=0}^{N} \Pr(state\_i) \times Loss(state\_i) \tag{4.23}$$

4.3 Heuristic Algorithms

Since solving an MILP problem is computationally intensive, heuristic algorithms are developed. SLNS-Heur (streamline effect based normal state route optimization heuristic) is developed to solve the NSR problem, while SLFR-Heur (streamline effect based failure recovery route optimization heuristic) is developed to solve the FRR problem.

4.3.1 Streamline Effect Based Normal State Route Optimization Heuristic (SLNS-Heur)

SLNS-Heur works in two steps. The first step performs initialization, where each flow is assigned the shortest path. The second step adopts iterative techniques to improve the route layout. In each iteration, some flows are randomly chosen to have their routes re-computed. If the new routes help reduce the overall burst loss, these flows will have their routes updated. Otherwise, the original routes are kept. The details of SLNS-Heur are described as below:

1. Each flow is assigned the path with the minimum cost. The cost of a link is the loss introduced by the new flow according to the loss estimation formula proposed. The details of the link cost calculation are given at the end of the algorithm description.
2. M flows are randomly selected and their traffic loads removed along the routes. Then the same link cost calculation method as that in step 1 is used to find a minimum-cost path for each of the M flows. If the re-routing reduces the overall burst loss, the routes for these M flows are updated. Otherwise the original routes are kept.
3. Step 2 is repeated until the stopping criterion is met.

The cost of link $< m, n >$, $cost(m, n)$, is computed as follows:

$$new = G(\rho_E^{m,n} + \rho', W) - G(\theta_E^{P(m),m,n} + \rho', W) \tag{4.24}$$

$$old = G(\rho_E^{m,n}, W) - G(\theta_E^{P(m),m,n}, W) \tag{4.25}$$

$$cost(m, n) = new - old + \delta \tag{4.26}$$

The notations used above are:

- ρ': the traffic load of the flow whose route is to be determined.
- $\rho_E^{m,n}$: the existing traffic load over link $< m, n >$.
- $P(m)$: the node prior to node m in the shortest path from the source node to node m.

- *new* and *old* : The difference between *new* and *old* is the increase in the burst loss over the link if the new flow is introduced. The burst loss over link $< m, n >$ with and without the flow going through are $G(\rho_E^{m,n} + \rho', W) - G(\theta_E^{P(m),m,n} + \rho', W) - \sum_{p \neq P(m)} G(\theta_E^{p,m,n}, W)$ and $G(\rho_E^{m,n}, W) - G(\theta_E^{P(m),m,n}, W) - \sum_{p \neq P(m)} G(\theta_E^{p,m,n}, W)$, respectively. Since we are concerned with the loss difference between these two values, we remove the common item of $- \sum_{p \neq P(m)} (\theta_E^{p,m,n}, W)$ and have the expression of *new* and *old* as given above.

Computational Complexity: Let the number of flows be F, the maximum node degree be Y, the number of nodes be V, the number of wavelengths per link be W, the total number of iterations be I, and the number of flows whose routes are re-computed each iteration be M. Since the complexity of Erlang B formula is $O(W)$, the complexity of one link cost computation is $O(Y \cdot W)$. The complexity of the shortest route search for each flow is $O(V^2)$. So the complexity of the first and the second steps are $O(F \cdot V^2 \cdot Y \cdot W)$ and $O(I \cdot M \cdot V^2 \cdot Y \cdot W)$, respectively.So the complexity of SLNS-Heur is $O(F \cdot V^2 \cdot Y \cdot W + I \cdot M \cdot V^2 \cdot Y \cdot W)$. We note that this complexity is reasonable and acceptable since the route optimization is done offline.

4.3.2 Streamline Effect Based Failure Recovery Route Optimization Heuristic (SLFR-Heur)

SLFR-Heur works in two phases. In the first phase, only the primary paths of all the flows are decided using SLNS-Heur, which is proposed above, to minimize the value of $Loss(state\_0)$. In the second phase, the primary paths are fixed, and we determine the backup paths to minimize the expected loss in failure states, EFL ($EFL = \sum_{i=1}^{N} (\frac{\Pr(state\_i)}{1-\Pr(state\_0)} \times Loss(state\_i))$.

It is acceptable to determine the primary/backup paths separately due to the following observations:

1. It can avoid negligence of backup path optimization. In reality, a communication network is in normal state mostly. Statistics show that, in most cases, 99.9% or more of the time, a network is working properly. Therefore, in the summation expression of the expected burst loss, the weighted loss in the normal state, $\Pr(state\_0) \times Loss(state\_0)$, has much more significance over the failure states. As a result, if we do not separate the optimizations of backup and primary paths, the heuristic algorithm might concentrate too much on the primary path optimization and neglect the backup path optimization.

2. It is known that determining the primary and backup paths separately may have the trap topology problem, which causes the backup path unavailable, though the network is two-connected. An example is given in Figure 4.2, the cost of each link being marked. Though there exist two link-disjoint paths between A and D, if the shortest route A->B->C->D is chosen as the primary path, there will be no backup paths between the node A and D.
   However, researches [67] have shown that such trap topologies are very rare in practice. More than 99.9% of the topology do not have this problem. So the chance is very small that the separate determination of primary/backup paths results in inefficient solutions.

3. Finding a shortest path is always simpler than finding a shortest pair of link-disjoint paths.

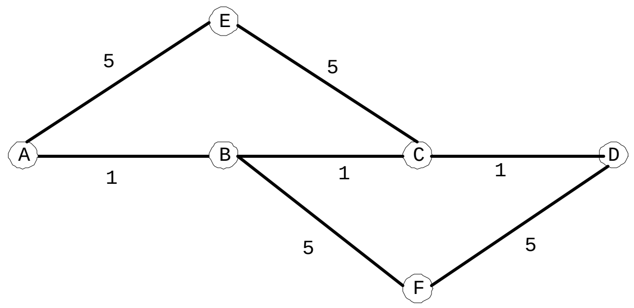

Figure 4.2: An example of trap topology problem

The description of first phase of SLFR-Heur is omitted, since it is the same as SLNS-Heur presented in the earlier section. The second phase of SLFR-Heur works as follows. First, each flow is assigned the shortest path link-disjoint to its primary path as the backup path. Then, iterative techniques are adopted to optimize the layout of backup routes. In each iteration, some flows are randomly chosen to have their backup routes re-computed. If the new routes help reduce the value of EFL, these flows will have their backup routes updated. In more details, the second phase of SLFR-Heur works as follows:

1. For each flow, assign the shortest path link-disjoint to its primary path as the backup path.

2. Randomly select M flows and remove their loads in the backup paths. Then for each of the M flows, find out the shortest path link-disjoint to the primary path as the backup path. The cost of a link is the increase in expected loss in failure states introduced by the new flow. The streamline effect loss estimation formula is used to estimate the loss. The details of the link cost definition are given later. After the backup path of a flow is decided, update the load of the links along the path. If the re-routing of backup paths reduces the expected burst loss in failure states, update the backup routes for these M flows. Otherwise keep the original

routes.

3. Repeat step 2 until the stopping criterion is met.

Suppose the flow whose route is to be determined has a traffic load of ρ_{new}. The cost of link from m to n, $cost(m, n)$, is computed as follows:

1. Calculate the value of $\rho_i^{m,n}$, the traffic load of all the flows going through each link $< m, n >$ in each state i, and the value of $\theta_i^{P(m),m,n}$,the traffic load going through both link $< P(m), m >$ and $< m, n >$ in each state i.

2. If the link is in the primary path, set the cost as infinity to refrain the backup path from traversing it. If the link is not on the primary path, estimate the expected loss increment brought by the new flow if this link is taken as the next hop in backup route searching. Note that, in the estimation, we can ignore those failure states where the failed link is not in the primary path and are only concerned with the failure states where the primary path is disrupted and the backup path is activated.

The pseudo-code of the computation is given below.

If (link$< m, n >$is on the primary path of the flow)

$cost(m, n) = \infty$;

else

{

$cost(m,n) = \delta;$

for (every failure state i)

{

if (link i is on the primary path)

{

$new = G(\rho_i^{m,n} + \rho', W) - G(\theta_i^{P(m),m,n} + \rho', W);$

$old = G(\rho_i^{m,n}, W) - G(\theta_i^{P(m),m,n}, W);$

$cost(m,n) += \frac{\Pr(state\_i)}{1 - \Pr(state\_0)} \times (new - old);$

}

}

}

Computational Complexity: We use the same notations as in the analysis of SLNS-Heur. In SLFR-Heur, the complexity of the first phase is the same as SLNS-Heur. The second phase is

N times more complex than SLFR-Heur, since the link cost computation involves N failure states. Therefore, the complexity of SLFR-Heur is $O(F \cdot V^2 \cdot Y \cdot W \cdot N + I \cdot M \cdot V^2 \cdot Y \cdot W \cdot N)$. We note that this complexity is reasonable and acceptable since the route optimization is done offline.

4.4 Numerical Results

In this section, the performance of routing algorithms proposed in this chapter are compared with other known algorithms. Experiments are conducted over three network topologies, a 10-node network shown in Figure 4.3, the 14-node NSFNET topology shown in Figure 3.5 and the 33-node Pan-European topology (see http://www.geant.net). In the Pan-European topology, some links are added to make the network bi-connected. The revised topology is shown in Figure 4.4, with the links added shown in dashed lines. Each link carries 32 wavelengths. Two traffic load scenarios are considered. The first one is the identical scenario, where the traffic load of each flow is equal. The second one is the non-identical scenario where the traffic load of each flow is uniformly randomly chosen between $0.5c$ and $1.5c$, where c is the average load of flows. The arrival of bursts in each flow follows Poisson distribution. In the experiments, after the route layouts are computed, we use the simulator developed by us to measure the burst loss rates of different route layouts.

4.4.1 Performance Study for the NSR Problem

The following algorithms are implemented to solve the NSR problem and the performance of the route layouts found by them is compared.

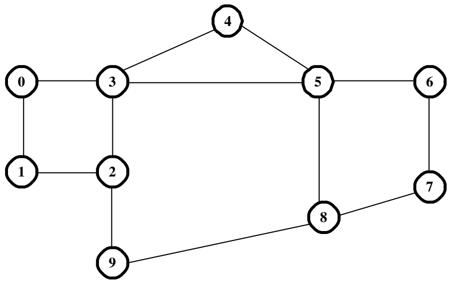

Figure 4.3: A 10-node network topology

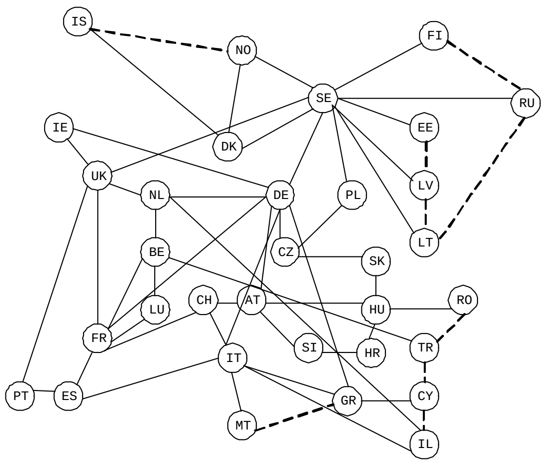

Figure 4.4: Pan-European Topology

1. MILP1 presented in this chapter to solve the NSR problem. The interpolation points of ρ in the piecewise function $\hat{G}(\rho, W)$ are 6.4, 12.8, 19.2 and 25.6. Note that the loads are the offered load. The value of corresponding normalized load can be derived by division over the number of wavelengths on a link.

2. SLNS-Heur presented in this chapter. The optimization process stops when the objective value reduction is less than 0.1% in the past 100 iterations or the number of iterations reaches 1000, whichever is satisfied first.

3. The Erlang B based route layout MILP presented in [6] (referred as Erl-MILP), where the streamline effect is not considered and Erlang B formula is used to estimate the loss. The parameters are set as in [6].

4. The Erlang B based route layout heuristic presented in [6] (referred as Erl-Heur), where the streamline effect is not considered and Erlang B formula is used to estimate the loss. The parameters are set as in [6].

5. The widest shortest path algorithm (referred as WSP). Here the flows are considered in a non-increasing order of their traffic load. We found that such an order achieves the best performance compared with the other routing order we have tested, including the random order, the least-load-first order, the longest-flow-first order and the shortest-flow-first order.

6. An MILP formulation to minimize the maximal load on any link (referred as Min-max).

7. The shortest path first algorithm (referred as SPF), which assigns the minimum-hop path to each flow.

Table 4.1: Burst Loss Rates in the 10-Node Network

No. of Flows	MILP1	SLNS-Heur	Erl-MILP	Erl-Heur	WSP	Min-max	SPF
8	2.71e-7	2.71e-7	6.79e-5	6.79e-5	2.03e-3	1.35e-4	7.79e-3
10	2.26e-6	2.83e-6	6.12e-4	6.12e-4	1.20e-2	1.27e-3	1.28e-2
12	5.10e-4	5.10e-4	2.03e-3	2.13e-3	3.78e-2	4.79e-3	5.90e-2

In the experiments on the 10-node network, all the algorithms are evaluated. In the experiments over the NSFNET and the Pan-European topologies, only four algorithms, SLNS-Heur, Erl-Heur, WSP and SPF, are evaluated, because the MILPs are computationally intensive in these larger networks.

4.4.1.1 Results for 10-Node Network

We consider three scenarios, where 10, 11 and 12 node pairs are chosen to be ingress-egress pairs, respectively. The load of each flow ρ =6.4. In SLNS-Heur, the routes for four flows are re-computed in each iteration. The results are listed in Table 4.1. As the results show, the MILP1 and SLFR-Heur methods give lower overall burst loss than the other methods. Thus the route layouts computed by our methods are better. We can see that SLNS-Heur achieves the same performance as MILP1 in two out of the three scenarios. It shows that SLNS-Heur can find route layouts close to optimal, at least in simple scenarios.

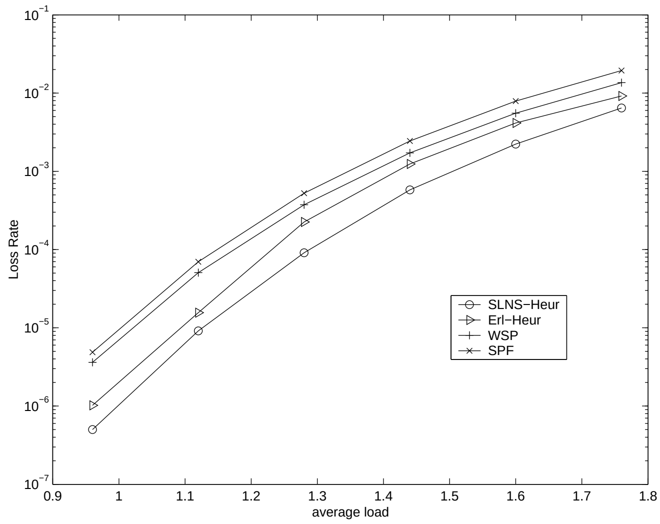

Figure 4.5: Burst Loss Rates of Different Algorithms (NSFNET topology, Identical Load Scenario)

4.4.1.2 Results for NSFNET Topology

There is a flow between each node pair, thus there are a total of 182 flows. Both the identical and the non-identical load scenarios are tested. In SLNS-Heur, the routes of 15 flows' are re-computed in each iteration. We observe that SLNS-Heur gives lower overall burst loss than the other methods for both the identical and the non-identical load scenario as shown in Figure 4.5 and Figure 4.6. We also observe that when the traffic load is light the performance gain of SLNS-Heur over the other methods is greater. The gain decreases as the load increases. We will discuss this phenomenon later.

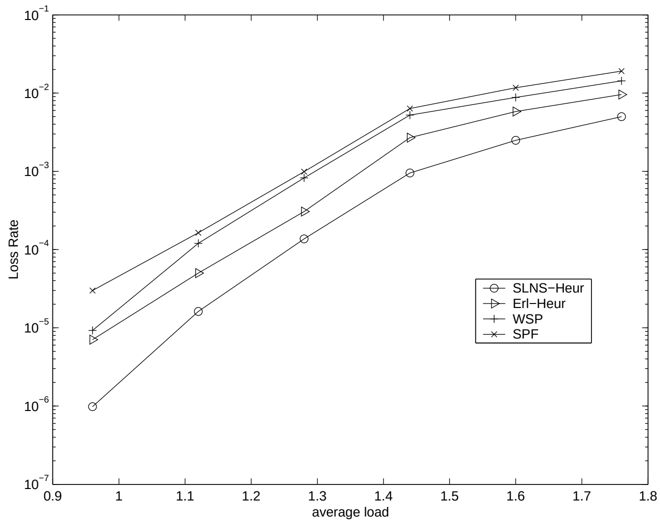

Figure 4.6: Burst Loss Rates of Different Algorithms (NSFNET topology, Non-Identical Load Scenario)

4.4.1.3 Results for Pan-European Topology

There is a flow between each node pair, and thus there are a total of 1056 flows. Both the identical and the non-identical load scenarios are tested. In SLNS-Heur, the routes for 15 flows are re-computed in each iteration. We see that SLNS-Heur gives lower overall burst loss than the other methods for both the identical and non-identical load scenarios as shown in Figure 4.7 and Figure 4.8. We also observe that when the traffic load is light the performance gain of SLNS-Heur over the other methods is greater. It is due to the following reason. In the Erlang B formula, the first derivative is positive and the second derivative is negative, so the loss rate increases more slowly when the load increases. Similarly, when estimating with the streamline effect formula, the loss rate will increase more slowly as the load increases. As a result, when the load is heavier, the difference caused by different route layouts are smaller. So the performances of the different algorithms tend

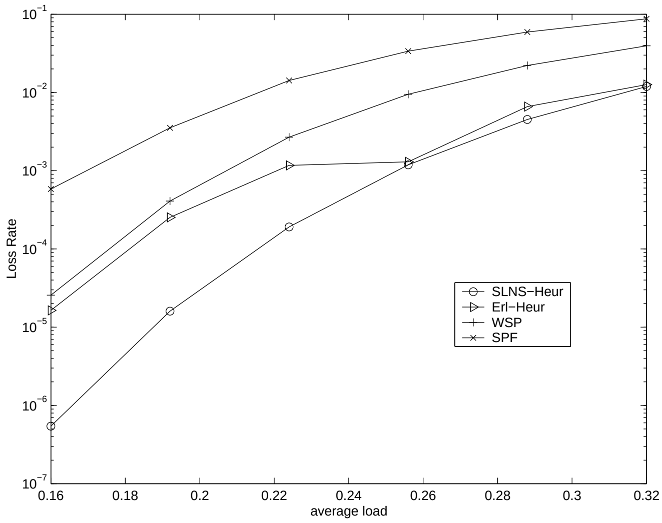

Figure 4.7: Burst Loss Rates of Different Algorithms (Pan-European Topology, Identical Load Scenario)

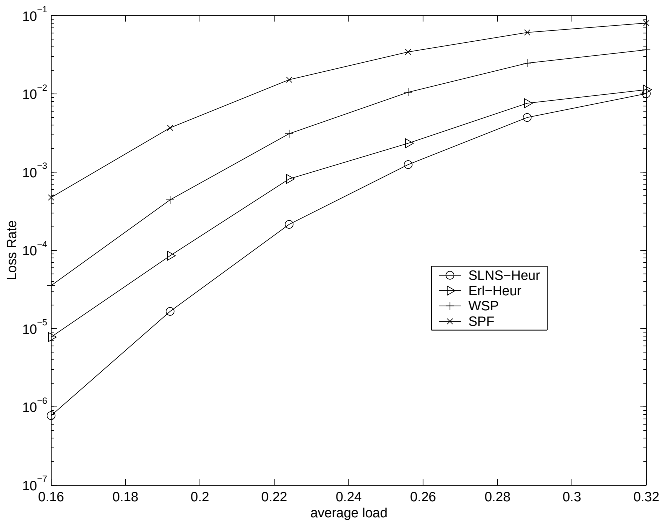

Figure 4.8: Burst Loss Rates of Different Algorithms (Pan-European Topology, Non-Identical Load Scenario)

to be closer. Such behavior has been observed by other researchers [6].

4.4.2 Performance Study for the FRR Problem

The following algorithms are implemented to solve the FRR problem and the performance of the route layouts found by them is compared.

- MILP2 proposed in this chapter to solve the FRR problem. The interpolation points of ρ in the piecewise function $\hat{G}(\rho, W)$ are 6.4, 12.8, 19.2 and 25.6.
- SLFR-Heur proposed in this chapter. In both phases, the optimization process stops when the objective value reduction is less than 0.1% in the past 100 iterations or the number of iterations reaches 1000, whichever is satisfied first.
- The Erlang B based failure recovery route layout MILP formulation extended from [6] (referred as Ext-Erl-MILP), where the loss estimation is based on the Erlang B formula. The parameters are set as in [6].
- The Erlang B based failure recovery route layout heuristic algorithm extended from [6] (referred as Ext-Erl-Heur). In Ext-Erl-Heur, first the heuristic in [6] is used to determine the primary paths. Then, an algorithm similar to the second phase of SLFR-Heur, except that the link cost estimation is based on the Erlang B formula, is used to determine the backup paths. The parameters of the first phase are set as in [6], and the parameters of the second phase are set the same as the second phase of SLFR-Heur.

- An MILP formulation to minimize the maximal load in any link in any states (referred as Min-max).
- The shortest path algorithms (referred as SPF), where the minimum-hop path is selected as the primary path, and the minimum-hop path disjoint to the primary path is chosen as the backup path.

In each topology, the probability of the normal state is 99.9% and the 0.1% failure probability is evenly distributed among all the links in the network. In the experiments over the 10-node network, all the algorithms are evaluated. In the experiments over the NSFNET and the Pan-European topologies, only three algorithms, SLFR-Heur, Ext-Erl-Heur, SPF, are evaluated, as the MILPs are computationally intensive for these larger networks.

4.4.2.1 Results for 10-Node Network

We consider three traffic scenarios, where 4, 5 and 6 node pairs are chosen to be ingress-egress pairs, respectively. The load of each flow ρ =6.4. In both phases of SLFR-Heur and the second phase of Ext-Erl-Heur, the routes for 3 flows are re-computed in each iteration. The results are listed in Table 4.2 and Table 4.3. As the results show, the proposed MILP2 and SLFR-Heur methods give lower expected burst loss over normal and failure states and expected burst loss in failure states than the other methods. Thus the primary route layouts and also the backup route layouts computed by our methods are better. We can see that SLFR-Heur achieves the same performance as MILP2 in two out of the three scenarios. It shows that SLNS-Heur can find route layouts close

Table 4.2: Expected Burst Loss Rate over Normal and Failure States in the 10-Node Network

No. of flows	MILP2	SLFR-Heur	Ext-Erl-MILP	Ext-Erl-Heur	Min-max	SPF
4	2.72e-10	2.72e-10	3.26e-10	3.26e-10	5.99e-10	1.33e-9
5	3.05e-10	3.05e-10	1.13e-7	1.13e-7	9.57e-7	1.13e-6
6	5.08e-10	6.38e-10	1.02e-6	8.61e-5	9.43e-5	2.08e-6

Table 4.3: Expected Burst Loss Rate in Failures in the 10-Node Network

No. of flows	MILP2	SLFR-Heur	Ext-Erl-MILP	Ext-Erl-Heur	Min-max	SPF
4	2.72e-7	2.72e-7	3.26e-7	3.26e-7	5.98e-7	1.31e-6
5	3.04e-7	3.04e-7	1.21e-6	1.21e-6	2.13e-6	1.71e-4
6	5.07e-7	6.37e-7	7.92e-5	1.02e-4	1.27e-4	1.97e-4

to optimal, at least in simple scenarios.

4.4.2.2 Results for NSFNET Topology

Each node pair is an ingress-egress pair, i.e., there are 182 flows. Both the identical and the non-identical load scenarios are tested. In both phases of SLFR-Heur and the second phase of Ext-Erl-Heur, the routes for 15 flows are re-computed in each iteration. It is observed that SLFR-Heur gives the lowest expected burst loss over normal and failure states for both the identical and the non-identical load as shown in Figure 4.9 and Figure 4.10. Figure 4.11 and Figure 4.12 show that

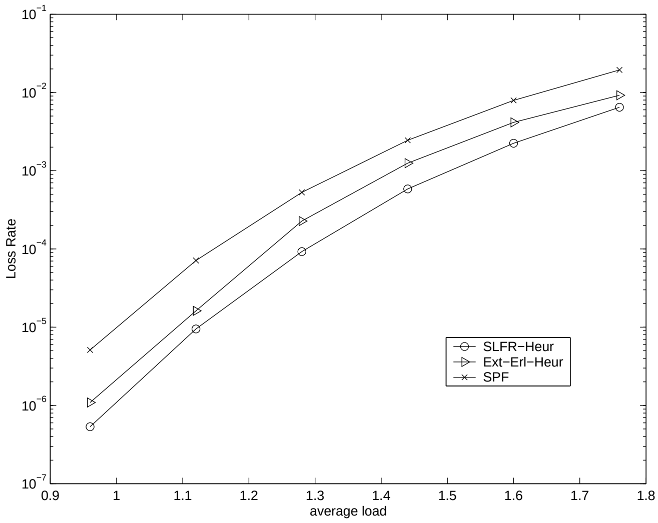

Figure 4.9: Expected Burst Loss Rates over Normal and Failure States of Different Algorithms (NSFNET, Identical Load Scenario)

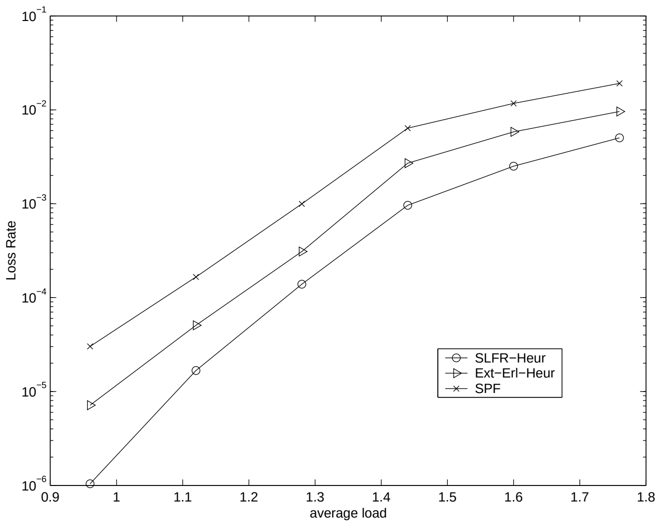

Figure 4.10: Expected Burst Loss Rates over Normal and Failure States of Different Algorithms (NSFNET, Non-Identical Load Scenario)

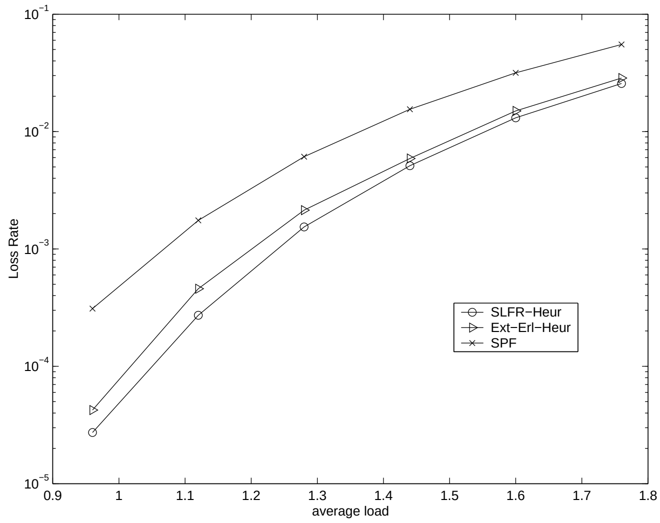

Figure 4.11: Expected Burst Loss Rates in Failure States of Different Algorithms (NSFNET, Identical Load Scenario)

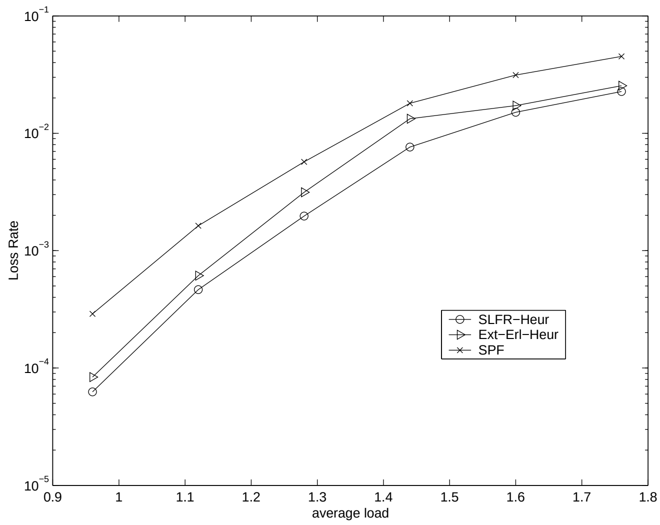

Figure 4.12: Expected Burst Loss Rates in Failure States of Different Algorithms (NSFNET, Non-Identical Load Scenario)

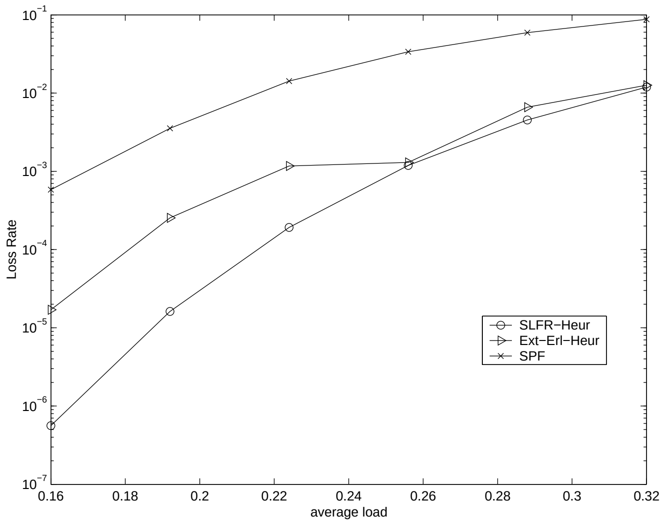

Figure 4.13: Expected Burst Loss Rates over Normal and Failure States of Different Algorithms (Pan-European, Identical Load Scenario)

SLFR-Heur gives the lowest expected burst loss in failure states for both identical and non-identical load scenarios. So both the primary and the backup route layouts determined by SLFR-Heur are better. We also observe that when the traffic load is light the performance gain of SLFR-Heur over the Ext-Erl-Heur is greater. Besides, the gain of our algorithm is less in terms of expected burst loss over failure states than in terms of the expected burst loss over normal and failure states. We will discuss these observations later.

4.4.2.3 Results for Pan-European Topology

Each node pair is an ingress-egress node, i.e., there are a total of 1056 flows. Both the identical and the non-identical load scenarios are tested. In both phases of SLFR-Heur and the second phase of Ext-Erl-Heur, the routes for 15 flows are re-computed in each iteration.

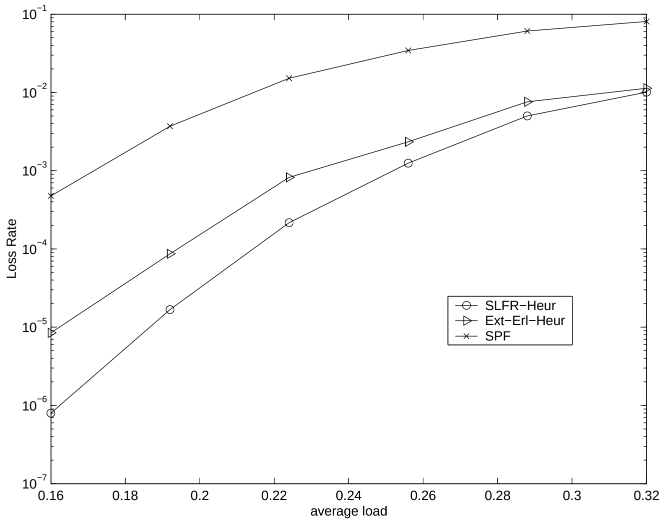

Figure 4.14: Expected Burst Loss Rates over Normal and Failure States of Different Algorithms (Pan-European, Non-Identical Load Scenario)

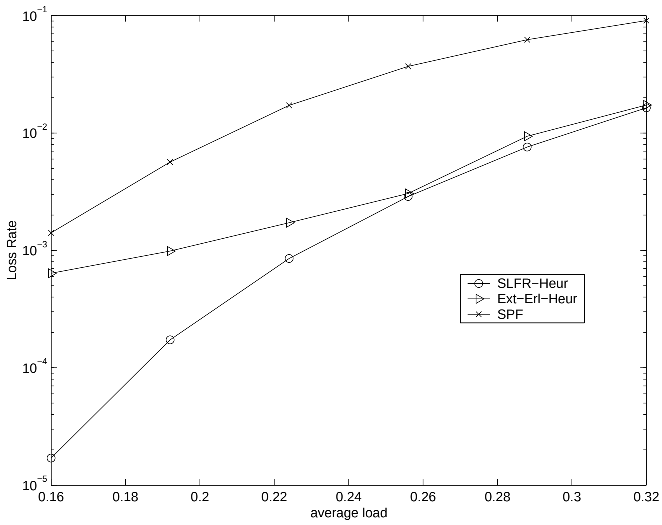

Figure 4.15: Expected Burst Loss Rates in Failure States of Different Algorithms (Pan-European, Identical Load Scenario)

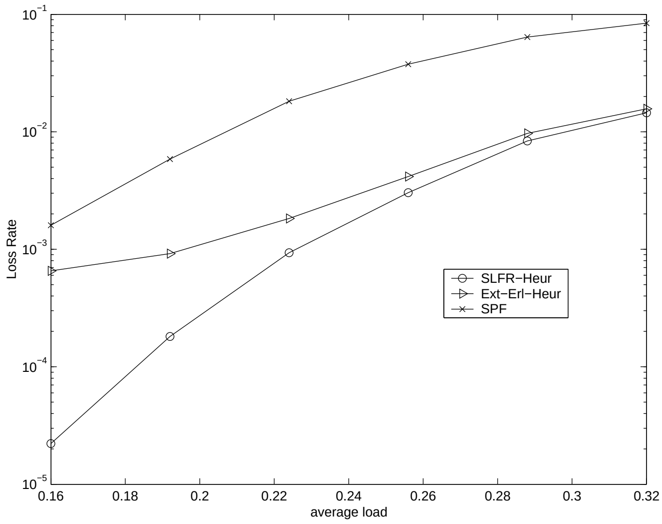

Figure 4.16: Expected Burst Loss Rates in Failure States of Different Algorithms (Pan-European, Non-Identical Load Scenario)

We see that SLFR-Heur gives the lowest expected burst loss over normal and failure states for both load scenarios in Figure 4.13 and Figure 4.14. Figure 4.15 and Figure 4.16 show that SLFR-Heur gives the lowest expected burst loss in failure states for both load scenarios. So both the primary and the backup route layouts determined by SLFR-Heur are better. We also observe that when the traffic load is light the performance gain of SLFR-Heur over Ext-Erl-Heur is greater. This phenomenon has been discussed earlier in the performance study for the NSR problem. Besides, the gain of our algorithm over the others in terms of the expected burst loss over failure states is less than in terms of the expected burst loss over normal and failure states. The reason is that when we choose two link-disjoint paths instead of one (i.e. the primary path only), the space to optimize the backup path is smaller. Similar observations were made in [6]. In [6], the performance enhancement is much less in primary-backup path selection than in primary route selection, though similar route selection techniques are used. However, the difference in this topology is not so large as in the case of the NSFNET topology. This is because the Pan-European network is larger and denser and there are more choices to select the backup path for a given primary path.

4.5 Summary

Two problems of offline route optimization in OBS networks have been studied in this chapter. The first problem is to determine a route for each flow to minimize the overall burst loss. The second problem considers the failure states and determines the primary and backup paths for each flow to minimize the expected burst loss over the normal and the failure states. We have shown that the route selection based on Erlang B Formula or load balancing is inadequate due to the ignorance of the streamline effect. Based on the new formula considering streamline effect, we have developed MILP formulations for the two problems. Since MILP-based solutions are computationally intensive, we have proposed heuristic algorithms to solve the two problems. The simulation results show that our algorithms are very effective in finding the route layouts that yield lower burst loss than other known algorithms.

Chapter 5

Residual Admission Capacity : A Metric to Measure Link Residual Capacity in OBS Networks

This chapter analyzes the unique feature of OBS networks in link residual capacity measurement. First we discuss the importance of estimating link residual capacity in configuring loss guaranteed connections. After that, we show that it is inaccurate to apply the same method of computing residual bandwidth in IP or ATM networks in OBS networks directly because of the bufferless core nodes in OBS networks. Then we present the definition of the metric of residual admission capacity, which measures the amount of residual capacity on a link in OBS networks more accurately. Then a method to compute RAC is presented.

5.1 Importance of Residual Capacity Estimation

The following three chapters discuss the problems of configuring loss guaranteed tunnels (LGTs) in OBS networks with multi-protocol label switching (MPLS) control. An LGT is a connection, i.e. label switching path (LSP), with an associated end-to-end loss guarantee. LGTs are usually long-lived. They are designed and created by service providers based on the estimated traffic demand. The IP flows towards the same egress node with a specific loss requirement can be mapped to an appropriate LGT at the ingress node provided the total offered load is no larger than the maximal permitted load (referred as 'load' in the rest of the paper) of the LGT. In each link, the loss rate of an LGT is guaranteed to be lower than a certain value so that the end-to-end loss requirements are met.

Estimation of link residual capacity is crucial for configuring LGTs in such a way to improve network capacity utilization. In IP or ATM networks, usually each node has a large buffer and the total effective bandwidth required to meet the QoS demands of all the connections over a link is approximately the summation of the effective bandwidth required by each connection. The effective bandwidth of a connection is the minimal bandwidth needed to support the connection's QoS requirements. To compute the amount of residual capacity, first the effective bandwidth of each connection over this link is computed, and the residual bandwidth is the total capacity of the link subtracted by the effective bandwidth of each connection. However, due to the bufferless feature, such a computation method is not applicable to OBS networks. The total resource required by the aggregated connections is no longer the summation of the resource required by each connection. If we adopt the residual bandwidth computation methods used for IP networks to OBS networks directly, we will obtain inaccurate results. As we will show later, the computation may show that the resource on a link is used up when there is still some capacity available to admit new connections. Also, it may show that a link has a larger residual capacity than another link where actually there is less resource available. It motivates us to define a new metric, called *residual admission capacity (RAC)*, which gives a more accurate description on how large the residual capacity is on a link in OBS networks.

5.2 Inaccuracy of Traditional Residual Bandwidth Computing Method in OBS Networks

The residual capacity on a link is measured by its ability to accommodate more loss guaranteed traffic. Suppose that there are two links. The first link can accommodate a new LGT without violating the loss requirements of all the existing LGTs and the new LGT, while the second link cannot. We say that the first link has sufficient residual capacity for this new LGT but not the second link. In IP or ATM networks, to compute the amount of residual capacity, first the effective bandwidth of each connection over this link is computed. Effective bandwidth is the minimal capacity needed to meet the QoS requirement of a connection. The residual bandwidth is the total capacity of the link subtracted by the effective bandwidth of each connection. We will show that this traditional method of computing residual bandwidth used in IP or ATM networks is not applicable in OBS networks.

As reviewed in Chapter 2, the preemptive scheme [9][10] shows higher efficiency than the other existing loss guarantee provisioning schemes. So in our work we suppose the preemptive scheme or a similar scheme is implemented in each node to provide loss guarantees to LGTs. In such a scheme, the total burst loss does not increase compared to the case wherein no loss guarantee provisioning mechanism is implemented, and the distance of actual loss rate of each LGT to the corresponding loss guarantee is kept almost equal.

We now discuss how to decide whether the loss guarantees of LGTs on a link can be satisfied. As an LGT in OBS networks is the aggregation of many independent IP streams, we assume that the bursts of each LGT follow Poisson arrival. Besides, we also assume that the nodes are not equipped with fiber delay lines (FDLs). Therefore, the burst loss rate on a link can be determined using the Erlang B formula as follows:

$$L(\rho, W) = \frac{(\frac{\rho^W}{W!})}{(\sum_{m=0}^{W} \frac{\rho^m}{m!})} \tag{5.1}$$

Here ρ is the offered load on the link and W is the number of wavelengths carried by this link. $\rho = \lambda/\mu$, where λ and $1/\mu$ are the arrival rate and the mean burst length, respectively.

Suppose that there are B LGTs on the link and let ρ_i and g_i denote the offered load and the loss guarantee of the ith LGT, respectively. The weighted average threshold is defined as:

$$T = \frac{(\sum_{i=1}^{B}(\rho_i g_i))}{(\sum_{i=1}^{B}\rho_i)} \tag{5.2}$$

To ensure that the loss guarantee of the LGTs on a link is not violated, the overall burst loss probability needs to be kept no larger than the weighted average threshold. i.e. the following inequality should be satisfied:

$$T \geq L(\sum_{i=1}^{B}\rho_i, W) \tag{5.3}$$

By using Eq(5.1) and Eq(5.2), the above inequality can be rewritten as

$$\sum_{i=1}^{B}(\rho_i g_i) \geq \frac{\frac{(\sum_{i=1}^{B}\rho_i)^W}{W!}}{\sum_{m=0}^{W}(\frac{(\sum_{i=1}^{B}\rho_i)^m}{m!})}\sum_{i=1}^{B}\rho_i \tag{5.4}$$

The above inequality can be used to check if a link has enough capacity to satisfy the loss guarantees of the carried LGTs.

In IP or ATM networks, usually each node is equipped with a large buffer. Under the assumption that each node has a sufficiently large buffer, the total effective bandwidth required to meet the QoS demands of all the connections over a link is approximately the sum of the effective bandwidth required by each connection. However, generally the nodes in OBS networks are equipped with no or very small buffers. As a result, the total capacity required by the aggregated LGTs is no longer the summation of the capacity required by individual LGTs. If we use the same method in IP or ATM networks, the results are inaccurate. We substantiate this using the following two examples. Let the effective bandwidth of an LGT be the minimum number of wavelengths needed to support its loss requirement and the residual bandwidth on a link be the total number of wavelengths subtracted by the effective bandwidth of each LGT.

Example 1: Consider a link with 8 wavelengths and two LGTs, whose offered loads are both 0.5 Erlang and loss guarantees are both 1%. According to Erlang B formula, for an LGT with an offered load of 0.5 Erlang, at least 4 wavelengths are needed to make the loss rate no larger than 1%. So the effective bandwidth of each LGT is 4. The sum of the effective bandwidth of the two LGTs is 8. It renders the residual bandwidth to be 0 and means there is no residual capacity left. This is incorrect because, according to Eq(5.4), only 5 wavelengths are needed to support the loss requirements of these two LGTs. Further, this link has enough residual capacity to support another 4 LGT requests whose offered loads are 0.5 Erlang and required loss guarantees are 1%.

Example 2: Consider two links each with 8 wavelengths. Suppose that on the first link there is only one LGT whose offered load is 1.5 Erlang and required loss guarantee is 1%. According to Erlang B formula, for an LGT with an offered load of 1.5 Erlang, at least 6 wavelengths are needed to make the loss rate no larger than 1%. So the effective bandwidth for this LGT is 6 and the residual bandwidth for this link is 2. Suppose that on the second link, there are two LGTs, whose loss thresholds are both 1% and offered loads are 0.4 Erlang and 0.5 Erlang, respectively. So, according to Erlang B formula, the effective bandwidths for these two LGTs are 3 and 4, respectively and the residual bandwidth on the second link is 1. The values of the residual bandwidth suggest that the first link has larger residual capacity. However, it is not true as explained next. Suppose that a new LGT request, whose offered load is 1.5 Erlang and required loss guarantee is 1%, needs to be routed on both links. The first link cannot admit this LGT request while the second link can according to Eq(5.4). Therefore the second link has larger residual capacity.

The failure of this traditional method is due to the ignorance of the multiplexing effect where there are large voids between bursts. Since there is no buffer at the core nodes in OBS networks, there are inevitably voids between the bursts of an LGT. When more than one LGT are multiplexed onto a link, the voids of an LGT can be used by other LGTs. Therefore, the number of wavelengths required by the aggregated LGTs is always smaller than the summation of the number of wavelengths required by each LGT alone. So the traditional method always overestimates the number of wavelengths needed by the aggregated LGTs. Such overestimation may also exist in the traditional IP or ATM networks [68]. However, in networks with sufficient buffers, the voids between the packets are much smaller, so the traditional residual bandwidth computation method gives quite acceptable results and is thus widely used. But in OBS networks, the voids are much larger and this method gives unacceptable results.

5.3 Residual Admission Capacity (RAC) in OBS Networks

We use a metric called residual admission capacity to measure the amount of remaining capacity on a link. The definition of residual admission capacity is as follows.

Suppose that a LGT request with the requirement of loss guarantee of α arrives at a link in an OBS network. Let β be the maximal load possible for this LGT request without violating the loss requirements of the existing LGTs and this LGT request. We define the **residual admission capacity (RAC)** of a link, denoted by $\overline{\beta}$, as the value of β that corresponds to admission of an LGT request with the loss guarantee of 0. If $\overline{\beta}$ is not less than 0, the loss requirements of the LGTs carried on the link can be supported and vice versa.

From the definition, we have the following equation

$$Loss(\overline{\beta} + \rho_e, W) = \rho_e \times T_e \tag{5.5}$$

where ρ_e is the load of the existing LGTs on this link, T_e is the existing weighted average threshold, and $Loss(\rho, W)$ is the formula to estimate the burst loss. If we assume that the arrival of each LGT follows Poisson arrival, $Loss(\rho, W) = \rho \times L(\rho, W)$ ($L(\rho, W)$ is defined in Eq(1)). However, it should be noted that it is not required that the bursts of LGTs follow Poisson arrival. More discussion on this will be given in Section 5.3.1.

Since a loss estimation formula takes a complicated form, it is hard to derive a close-form expression of $\overline{\beta}$. We suggest to use the following method to compute $\overline{\beta}$:

1. Use the bisection method to solve the equation $Loss(x, W) = \rho_e \times T_e$. Denote the root as $x = \iota$. Since the burst loss is a monotonically increasing function of the load, there exists one and only one solution for the above equation and the bisection method is sure to find the

root.

2. The value of RAC on this link, $\overline{\beta} = \iota - \rho_e$.

We define that the RAC of a link without any load to be $+\infty$.We can see that, according to Eq(5.5), when $\rho_e = 0$, $\overline{\beta}$ is also zero. It means that an empty link cannot admit new LGTs, which is not true. Therefore we define that the RAC of an empty link is $+\infty$.

For the two examples given in Section 5.2, we show that RAC gives a correct estimation of the amount of residual capacity on the links. In Example 1, $Loss(\overline{\beta} + 1, 8) = 1 \times 0.01$, so $\overline{\beta}$ =1.6. So there is still capacity available on this link to admit new LGTs. In Example 2, on the first link, $Loss(\overline{\beta} + 1.5, 8) = 1.5 \times 0.01$, so $\overline{\beta}$ =1.27, while on the second link, $Loss(\overline{\beta} + 0.9, 8) = 0.9 \times 0.01$, so $\overline{\beta}$ =1.6. Therefore the second link has more residual capacity.

5.3.1 Discussion on Other Traffic Models and Node Configurations

In the above discussion, we assumed that the burst arrivals of an LGT follow Poisson distribution and the Erlang B formula is used to estimate the loss. This assumption is widely used in the OBS research community due to its simplicity and acceptable effectiveness. However, there are other traffic models and loss estimation methods proposed in the literature. For example, in [69], the effect of burst assembly algorithms on the traffic was analyzed and a loss model based on that analysis was proposed. In [70], the correlation among different links is considered and a loss estimation method was proposed accordingly. Besides, there are a lot of works on system modelling and performance analysis [71][72][73][74][75][76][77][78][79][80]. These models can also work with the concept of RAC. The only difference is that, in Equation 5.5, $Loss()$ is the loss estimation formula proposed in the respective literature. Any traffic model can fit into this definition of RAC as long as the loss estimation formula is known. For example, if we use the streamline effect based loss estimation formula presented in Chapter 3, we have

$$Loss() = \rho \frac{(\frac{\rho^W}{W!})}{(\sum_{m=0}^{W} \frac{\rho^m}{m!})} - \sum (\rho_j \frac{(\frac{\rho_j^W}{W!})}{(\sum_{m=0}^{W} \frac{\rho_j^m}{m!})})$$

Here the offered load from the jth upstream link is given by ρ_j. The total offered load is given by $\rho = \sum \rho_j$. With a more accurate loss estimation formula, the estimation on the link residual capacity will be more accurate accordingly.

Further, in the above discussion we assumed that each node is not equipped with optical buffers, i.e. FDLs. However, the definition of RAC is also applicable with links equipped with FDLs. In [81], the loss estimation formula for links equipped with FDLs is given. Using that formula as the $Loss()$ in Equation 5.5, we can compute the RAC on links with FDLs.

In brief, though we use the Poisson arrival assumption and Erlang B loss estimation formula in this thesis, the definition of RAC is more versatile. It can work with different traffic models and node configurations. However, due to the simplicity of Poisson arrival assumption, we adopt it in the rest of this thesis.

5.4 Summary

In this chapter, we showed that the traditional method of computing residual bandwidth in IP or ATM networks is not applicable to OBS networks due to the bufferless nature of OBS networks. We introduced a new metric, called residual admission capacity (RAC). We showed that our metric gives a more accurate estimation on the amount of residual capacity on links in OBS networks.

Chapter 6

RAC Based Loss Budget Partitioning and Loss Threshold Selection for Loss Guarantee Tunnels

Based on the concept of RAC presented in last chapter, in this chapter, we propose algorithms to solve two problems in configuring loss guaranteed tunnels:

1. the loss budget partitioning problem. This problem studies how to determine the loss guarantee values of an OBS connection on the traversing links so that the end-to-end loss requirements are met and the network capacity utilization efficiency is maximized.
2. the loss threshold selection problem. This problem is to determine the set of possible loss

guarantee values, i.e. the loss thresholds, on the links in such a way that the capacity utilization efficiency is maximized. Due to scalability reasons, the possible loss guarantees on the links can only take a finite number of values. As a result, it is quite likely that the end-to-end loss guarantee granted is not equal to, but smaller than, required. Improper loss threshold setting will cause unnecessary capacity consumption and reduces the network utilization efficiency.

For the first problem, we present an RAC based loss budget partitioning (RAC-LBP) algorithm. Initially, RAC-LBP sets the loss guarantee on each link to the smallest loss threshold that the residual capacity permits. Then it recursively relaxes the loss guarantee on the link with the least RAC left by setting it to the next larger loss threshold, until right before the end-to-end loss guarantee will be violated. RAC-LBP has no constraints on the loss threshold setting, and it can effectively utilize the network capacity because it allows an LGT to take different loss guarantee values according to the amount of residual capacity on the traversing links.

We also develop an RAC based loss threshold selection (RAC-LTS) algorithm for the second problem. We assume that the traffic profile of the LGT requests are known. RAC-LTS is comprised of two phases. In Phase I, loss guarantees of LGTs on traversing link can take any positive real value instead of only a finite number of values. Then for each LGT request considered, we determine the loss guarantees over each traversing link in such a way that the minimal RAC value along the path after the LGT being deployed is maximized, aiming to defer the formation of bottleneck links as much as possible. In Phase II, a finite set of loss thresholds are selected to represent the loss guarantee values found in Phase I. The set of loss thresholds having the minimal overall quantization cost is selected as the solution. Experiment results show that RAC-LTS utilizes the network capacity more effectively. Using the loss thresholds selected by RAC-LTS, more LGTs can be established.

6.1 RAC Based Loss Budget partitioning (RAC-LBP) Algorithm

This section presents the RAC Based loss budget partitioning (RAC-LBP) Algorithm. We suppose the path for the LGT under study has been determined, and the loss guarantee on each traversing link is decided in such a way that the end-to-end loss requirements are satisfied and the minimal RAC on the links along the path after the LGT is deployed is maximized.

The problem under study can be stated as below: Given K loss thresholds, $H = (h_1, h_2... h_K)$, where $0 < h_1 < h_2... < h_K < 1$, on each link. Consider an LGT request whose path is $(p_1, p_2...p_D)$, where p_m is the mth link in the path. The load of the LGT is ρ' and it requires that the end-to-end loss requirements be no larger than γ. It is required to determine:

1. whether the path has enough capacity to accommodate this LGT. In other words, it is to determine whether there exists a set of loss thresholds on the links along the path, denoted as $(g_1, g_2...g_D)$, where $g_m \in H$, such that:
   - $\sum_{m=1}^{D} g_m \leq \gamma$, i.e. the end-to-end loss requirement of this LGT is satisfied.
     Note that the exact relationship between g_m and γ should be $1 - \prod_{m=1}^{D}(1 - g_m) \leq \gamma$. However, the approximate expression $\sum_{m=1}^{D} g_m \leq \gamma$ is simple and yet feasible in our study.

More explanation is given in the appendix A.

- on each link p_m along the path, the loss guarantees of both the existing LGTs and the new LGT are satisfied after accommodating the new LGT.

2. If this path has enough capacity to accommodate this new LGT, it is required to determine the set of loss thresholds which maximizes the minimum of $\overline{\beta_m}$ ($\overline{\beta}$ in link p_m) on the links along the path after the LGT is admitted.

RAC-LBP works in two phases. Phase I determines whether the LGT request can be admitted. It works as the following:

1. In each traversing link p_m, set g_m to the smallest loss threshold the link residual capacity permits. We solve the equation $Loss(\rho'+\rho_m^e, W) = \rho' \times x + \rho_m^e \times T_m^e$. ρ_m^e and T_m^e are the existing load and the weighted average threshold on link v_m, respectively. Then among the thresholds not larger than x, set g_m to the one closest to x . In other words, $g_m = h_k$ ($h_k \leq x < h_{k+1}$). If there is no such a loss threshold, the request is rejected and the algorithm terminates, otherwise, go to step 2.

2. If $\sum_{m=1}^{D} g_m \leq \gamma$, this path has enough residual capacity to accommodate this LGT and the algorithm continue to Phase II. Otherwise, the request is rejected and the algorithm terminates.

Phase II of RAC-LBP determines the set of loss guarantee values which maximize the minimal RAC of the traversing links after the LGT is admitted. To achieve this, RAC-LBP recursively relaxes the loss guarantee on the link with the minimal RAC until no such relaxation is possible. In details, it works as below:

1. With the new LGT request taking the loss guarantees determined so far, compute the RAC value of each traversing link. Find out the link v_j which has the minimal RAC. Set g_j to the next larger loss threshold. If the existing $g_j = h_K$, the largest possible loss threshold, and cannot be relaxed any more, we relax the loss threshold on the second-minimal-RAC link, and so on.

2. Repeat Step 1 until right before $\sum_{m=1}^{D} g_m > \gamma$, or all the values of g_m are equal to the largest possible loss threshold and thus cannot be increased any more.

Compared with the existing loss budget partitioning algorithms, RAC-LBP is flexible and efficient. The algorithms in [8][11] require an LGT to use the same loss guarantees along the path regardless the amount of residual capacity available on the links. Comparing to them, RAC-LBP utilizes the network capacity more effectively. The algorithm in [9][10] permits an LGT to take loss guarantees on the traversing links, but it demands the loss thresholds follow a logarithmically uniform distribution. RAC-LBP places no constraints on the loss threshold setting and is thus more flexible.

6.2 RAC Based Loss Threshold Selection (RAC-LTS) Algorithm

This section describes the RAC Based Loss Threshold Selection (RAC-LTS) Algorithm. As the loss guarantee for an LGT on each link can only take a finite number of values due to scalability problem, an LGT may not find loss guarantees on the traversing links which makes the end-to-end loss rate exactly equal to its requirement. In order not to violate user's requirements, the network operator has to offer an end-to-end loss rate smaller than demanded and some extra capacity has to be consumed. If the loss thresholds are not selected properly, such extra consumption would be large. It motivates us to study the problem of loss threshold selection, to determine K loss thresholds, $H = (h_1, h_2 ... \ h_K)$, where $0 < h_1 < h_2 ... < h_K < 1$, on each link so that the the network capacity utilization is maximized.

We assume that the profile of the loss guaranteed traffic is known. The information includes the approximate number of LGT requests, the probability distribution of the source-destination node pairs, the range and probability distribution of LGTs' load and end-to-end loss requirements. Note that it is the approximate information instead of the details of each LGT that is required. For each LGT, we assume the shortest path is taken.

Intuitively, we know that a smaller granularity in the loss threshold setting (namely, more loss thresholds) gives better performance assuming the same loss threshold selection technique is used. The scheme where the loss guarantee value are continuous can be considered as a special case where the number of loss thresholds is infinite, and it should be the optimal limit for any setting with a finite number of loss thresholds. Based on this observation, RAC-LTS solves the loss threshold selection problem in two phases. Phase I is the process of continuous loss threshold searching. In Phase I, the constraint on the finity of the loss threshold number is temporarily removed and loss thresholds can take any positive real value. We determine the loss guarantees for each LGT in each traversing link in such a way that the minimal RAC along the path after the LGT being deployed is maximized. Phase II is the process of loss threshold quantization. In Phase II, we determine K loss thresholds which minimize the extra resource consumption due to the constraint on the finity of the loss threshold number. Details of the two phases are given below.

6.2.1 Phase I: Continuous Loss Guarantee Searching

First, a batch of LGT requests, is generated according to the traffic profile given. Then the LGT requests are considered one by one in a random sequence. We determine the loss guarantees of LGTs on each traversing link. Consider an LGT, whose path is $(p_1, p_2 ... p_D)$, where p_m is the mth traversing link along the path. The load of this LGT request is ρ' and its end-to-end loss rate should not be larger than γ. It is required to determine the optimal set of loss guarantees for this LGT, $(g_1, g_2 ... g_D)$, which not only satisfies its end-to-end loss demand but also maximize the minimal RAC of the traversing links after the LGT is admitted. Note that this problem is similar to that studied in the RAC-LBP algorithm, but here the loss guarantee value can take any positive real value. So the RAC-LBP algorithm is not applicable.

To solve the problem, first we determine whether the path has enough capacity to admit this LGT. In each traversing link p_m, we determine the smallest loss guarantee this link can provide by solving the equation $Loss(\rho' + \rho^e_m, W) = \rho' \times x + \rho^e_m \times T^e_m$ and set $g_m = x$. If $\sum_{m=1}^{D} g_m \leq \gamma$, this path can admit this LGT request, otherwise this request is rejected.

If this path has enough residual capacity for this LGT request, we continue to determine the optimal loss guarantees on each traversing link, which satisfies the following conditions:

a) $\sum_{m=1}^{D} g_m \leq \gamma$

b). the minimum of $\overline{\beta_m}$ ($\overline{\beta}$ in link p_m) on the links along the path is maximized after the LGT being admitted.

We describe below the procedure to compute optimal loss guarantees.

Looking into constraint a), we can see that $\sum_{m=1}^{D} g_m = \gamma$ when the least capacity is consumed and the minimum of $\overline{\beta_i}$ is maximized. So we replace constraint a) as $\sum_{m=1}^{D} g_m = \gamma$. As to constraint b), we can see that the minimal $\overline{\beta_i}$ is maximized when all the $\overline{\beta_i}$ are equal. To understand this, we can imagine, if there are there are two links, saying link p_u and p_q, are not equal in RAC, and$\overline{\beta_u} > \overline{\beta_q}$. Then the smaller of the two, $\overline{\beta_q}$, can become larger by decreasing the loss guarantee in link p_u, g_u, and increasing g_q. But if all the $\overline{\beta}$ on the traversing links are equal, no such adjustments are possible.

Based on the above observations, we use the following procedure to compute g_m.

1. We construct the following equation system for the LGT under study:

$$Loss(\rho' + \rho_1 + \overline{\beta}', W) = \rho' \times g_1 + \rho_1 \times T_1 \tag{6.1}$$

$$\begin{aligned} & Loss(\rho' + \rho_2 + \overline{\beta}', W) = \rho' \times g_2 + \rho_2 \times T_2 \\ & \ldots\ldots \\ & Loss(\rho' + \rho_N + \overline{\beta}', W) = \rho' \times g_D + \rho_D \times T_D \\ & \sum_{m=1}^{D} g_m = \gamma \end{aligned} \tag{6.2}$$

Where ρ' is the load of the new LGT, ρ_i and T_i are the existing load and the weighted average threshold on link i, $\overline{\beta}'$ is the RAC value of each link after the new LGT is deployed. Note that when $\overline{\beta}'$ is larger, g_m is larger and so is the value of $\sum_{m=1}^{D} g_m$. So there exists one and only one value of $\overline{\beta}'$ to satisfy the above equation system and we can use the bisection method to find it.

After the value of $\overline{\beta}'$ is determined, compute $g_m(m = 1, ..., D)$ using the first D equations.

2. Check the validity of solution. As the loss guarantee can never be negative, if $g_m < 0$, set $g_m = 0$ and remove the mth equation to construct a new equation system. For example, in a 3-hop path, if $g_3 < 0$, we set $g_3 = 0$ and remove the third equation, then the new equation system is

$$\begin{aligned} & Loss(\rho' + \rho_1 + \overline{\beta}', W) = \rho' \times g_1 + \rho_1 \times T_1 \\ & Loss(\rho' + \rho_2 + \overline{\beta}', W) = \rho' \times g_2 + \rho_2 \times T_2 \\ & g_1 + g_2 = \gamma \end{aligned} \tag{6.3}$$

Solve the new equation system using the same techniques as in step 1. Then go back to step 2 to check the validity of the solution until all the $g_m \geq 0$.

In the same manner, we determine the continuous loss guarantees for each LGT.

6.2.2 Phase II: Loss Threshold Quantization

In Phase II, we determine K loss thresholds which minimize the extra resource consumption due to the constraint on the number of the loss thresholds. Denote the loss guarantees determined in Phase I for the nth accepted LGT as $(g_{n,1}, g_{n,2}...)$. Suppose with a set of finite loss thresholds, $H = h_1, h_2...h_K$, the set of loss guarantees for this LGT which maximize the minimal RAC along the path is $(q_{n,1}, q_{n,2}...)$. We define a metric, "overall quantization cost", σ, to describe the degree of extra capacity consumption due to the constraints on the finite loss threshold values. The optimal set of loss thresholds will give the minimal σ.

$$\sigma = \sum_{LGT\_n} (\sum_{hop\_i} (g_{n,i} - q_{n,i})) \tag{6.4}$$

We can see that $\sum\limits_{hop\_i} g_{n,i} \geq \sum\limits_{hop\_i} q_{n,i}$. It is due to that in continuous loss guarantee searching, $\sum\limits_{hop\_i} g_{n,i} = \gamma_n$ (γ_n is the end-to-end loss requirement of the nth accepted LGT), while with the finite loss threshold setting, $\sum\limits_{hop\_i} g_{n,i} \leq \gamma_n$. Since both $(g_{n,1}, g_{n,2}...)$ and $(q_{n,1}, q_{n,2}...)$ are determined in such a way that a larger guarantee is assigned at a link with less RAC and vice versa, we can expect that $q_{n,i}$ is very likely to be the maximal threshold which is no larger than $g_{n,i}$. Mathematically, it is $q_{n,i} = \max(h_j | h_j \leq g_{n,i}, h_j \in H)$. Note that the relationship need not hold always, but it occurs with a high probability.

Based on this observation, the above problem is reducible to the well-studied p-facility location problem in graph theory. P-facility problem can be described as follows: the sites of no more than P facilities need to be selected to serve a set of customers at different points. There is a fixed cost associated with setting up a facility and a cost of serving each customer from a given facility site. The objective is to minimize the cost of locating facilities and serving customer.

Our problem can be regarded as a P-facility location problem with the following reduction:

1. a loss threshold h_i is considered as a facility and the number of facilities, P, is equal to K in our problem.
2. each $g_{n,i}$ is considered as a customer.
3. the cost associated with setting up a facility is zero.
4. the cost of customer $g_{n,j}$ being served by facility h_i , $C_{n,ij}$ is defined as follows:

$$\begin{aligned} C_{n,ij} &= g_{n,j} - h_i, \quad \text{if} \;\; g_{n,j} \geq h_i \\ &= +\infty, \qquad\quad \text{if} \;\; g_{n,j} < h_i \end{aligned} \tag{6.5}$$

The general P-facility location problem is an NP-hard problem [82]. However, our problem is a special case of P-facility problem, where all the customers are on a line, and there exist polynomial algorithms for solving it. We use this algorithm proposed in [83] to determine the values of K loss thresholds.

6.3 Numerical Results

In this section, we study the performance of the proposed algorithms. We consider the 14-node NSFNET network and the 33-node Pan-European topology. As the results for both networks show similar performance trend, only the results for the NSFNET network are presented. Each link carries 32 wavelengths. We suppose that the network supports 5 end-to-end loss guarantee classes, numbered from 1 to 5, with the end-to-end loss guarantees equal to 1×10^{-4}, 5×10^{-4}, 1×10^{-3}, 5×10^{-3} and 1×10^{-2}, respectively. Between each node pair, there are 5 LGT requests, corresponding to the 5 classes. The LGTs are considered one by one in a random sequence. Each LGT uses the shortest path. Once an LGT is accepted, it stays there. We vary the average maximum permissible load (referred as "average load" in the rest of the chapter) of the LGT requests. For a given average load c, the load of each LGT is uniformly distributed between $1.5c$ Erlang and $0.5c$ Erlang. For each average load value, we generate 20 sets of LGT requests randomly and present the average of 20 trials.

6.3.1 Performance of Loss Budget partitioning Algorithms

We study the performance of our algorithm, RAC-LBP, in this section. We consider two different loss threshold settings, RAC-LBP(1) and RAC-LBP(2). In RAC-LBP(1), the loss thresholds are the same as in Log-LBP. In RAC-LBP(2), the loss thresholds are decided by RAC-LTS. We also compare RAC-LBP with the following algorithms.

1. Avg-LBP: the average loss budget partitioning scheme proposed in [8].

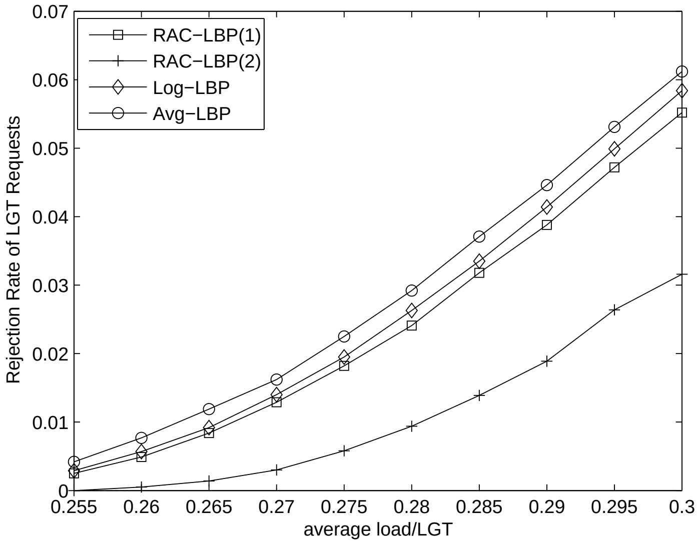

Figure 6.1: Rejection Rate of LGT Requests under Different Loss Budget Partitioning Algorithms

2. Log-LBP: the scheme proposed in [10].

   In Avg-LBP, the possible number of loss guarantee values on each link is equal to the number of end-to-end loss classes, which is 5 in our study. The number of loss thresholds in Log-LBP is also set to 5. The 5 loss thresholds, $h_1, \alpha h_1 ..., \alpha^4 h_1$, are are set in such a way that $h_1 = B^{e2e}_{\min}/Z$ and $\alpha^5 h_1 = B^{e2e}_{\max}$, where Z is the network diameter, $B^{e2e}_{\min}$ and $B^{e2e}_{\max}$ are the minimal and the maximal end-to-end loss guarantees provided.

Figure 6.1 compares the algorithms in terms of the average rejection rates. Figure 6.2 shows the relationship of rejection rate and path length. Figure 6.3 shows the rejection rates of different end-to-end loss classes. As the results show, RAC-LBP gives the lowest rejection rate. When using the same loss threshold setting as Log-LBP, RAC-LBP's performance gain over Log-LBP is small. However, we note that Log-LBP only works when the loss thresholds follow a uniform distribution

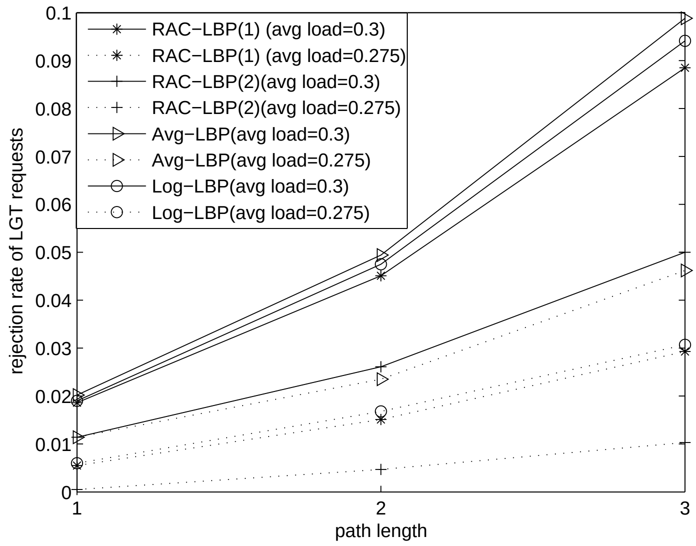

Figure 6.2: Rejection Rate of LGT Requests of Different Path Length under Different Loss Budget Partitioning Algorithms

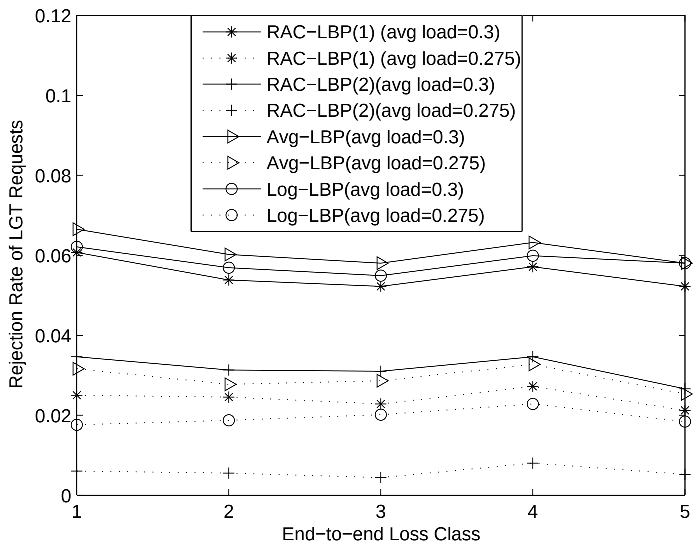

Figure 6.3: Rejection Rate of LGT Requests of Different End-to-End Classes under Different Loss Budget Partitioning Algorithms

on a logarithmic scale. With the loss thresholds set by RAC-LTS, RAC-LBP shows significant performance gain. Figure 6.2 shows that LGTs with longer paths have a higher rejection rate. This is because, longer-hop paths have a higher chance of encountering a bottleneck link with low residual capacity. Figure 6.3 shows that the rejection rates of different end-to-end loss classes are close, with the classes demanding smaller end-to-end loss rates showing slightly higher rejection rates. The reason can be explained as follows. The impact of hop length applies to all classes and the hop length effect is more pronounced than the variation in loss requirements by different classes.

6.3.2 Performance of Loss Threshold Selection Algorithms

We study the performance of our algorithm, RAC-LTS. We generate a set of LGT requests by random seed 0 to determine the loss threshold values, which are fixed in the 20 trials. We also compare RAC-LTS with the following algorithms.

1. Log-LTS: the logarithmic uniformly distributed setting proposed in [10]. The K loss thresholds $h_1, \varrho h_1 ..., \varrho^{K-1} h_1$ are set in such a way that $h_1 = B^{e2e}_{\min}/Z$ and $\varrho^K h_1 = B^{e2e}_{\max}$.
2. Uni-LTS: the uniformly distributed setting. The K loss thresholds $h_1, h_1 + d ..., h_1 + (K-1)d$ are set in such a way that $h_1 = B^{e2e}_{\min}/Z$ and $h_1 + Kd = B^{e2e}_{\max}$.
3. RAC-LTS-Inf: the loss guarantee searching technique in RAC-LTS Phase I. It is different from the others as the loss guarantee values are continuous.

For the algorithms of RAC-LTS, Log-LTS and Uni-LTS, we study the performance for two different numbers of loss thresholds, 2 and 4. Figures 6.4, 6.5 and 6.6 study the scenario where the network cannot accommodate all the requests and rejection happens. Figure 6.4 shows the rejection rates of different algorithms. Figure 6.5 shows the rejection rates of LGTs with different path lengths. Figure 6.6 shows the rejection rates of different end-to-end loss classes. Figures 6.7 and 6.8 show the amount of minimal RAC and average RAC of the links in the scenario where no requests are rejected.

We define a metric called Loss Guarantee Deviation Index (LGDI) to measure how close the loss guarantee provided is to the guarantee required. LGDI is defined as $\frac{\sum \Upsilon_i - \sum \Psi_i}{\sum \Psi_i}$, where Υ_i and Ψ_i are the end-to-end loss guarantees that the i^{th} accepted LGT request required and actually provided, respectively. It is desirable to keep this value as small as possible to avoid excessive resource usage leading to a lower rejection rate. The performance comparison of the algorithms in terms of LGDI is given in Figure 6.9.

We can make the following observations from the above results:

1. Of the three finite loss threshold selection algorithms, RAC-LTS gives the best performance. Its rejection rate is the minimal when rejection happens. Its average RAC and minimal RAC is the largest when there is no rejection. Its LGDI is the smallest. With only 2 thresholds, RAC-LTS achieves a performance better than Uni-LTS and Log-LTS with 4 thresholds.

2. Given the same algorithm, the more the number of loss thresholds, the better the performance is. In RAC-LTS-Inf, where the number of possible loss threshold values is infinite, the per-

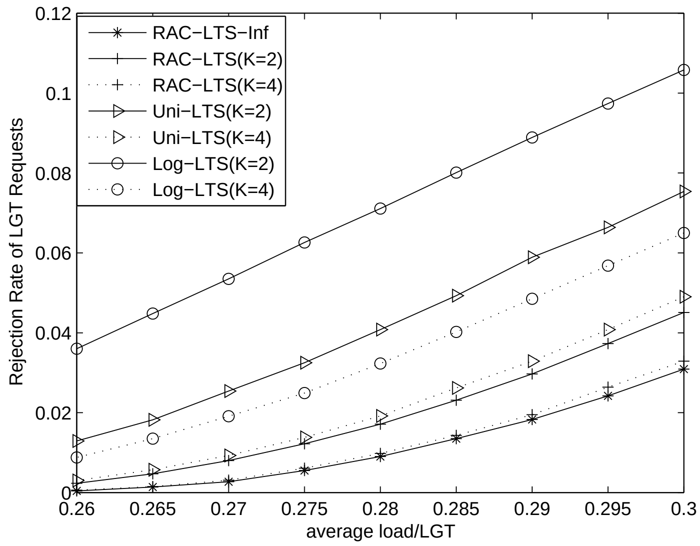

Figure 6.4: Rejection Rate of LGT Requests under Different Loss Threshold Selection Algorithms

formance is the best. It justifies the use of loss guarantee values determined by RAC-LTS-Inf to select the finite loss thresholds in RAC-LTS.

3. LGTs with longer paths have a higher rejection rate. However, the rejection rates of different end-to-end loss classes are close. The analysis for the same phenomenon for the loss budget partitioning problem is also applicable here.

4. In RAC-LTS, we use the estimated information of the LGT requests to determine the loss thresholds. We can expect that RAC-LTS will give good performance when there are minor changes in the loss-guaranteed traffic demand.

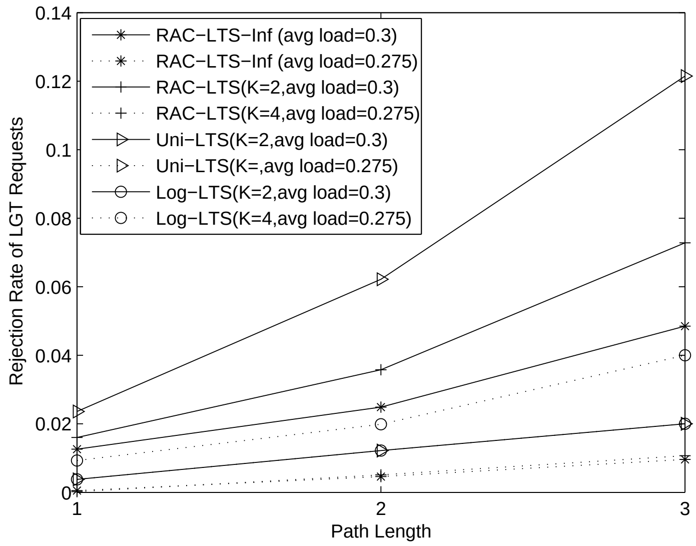

Figure 6.5: Rejection Rate of LGT Requests of Different Path Lengths under Different Loss Selection Algorithms

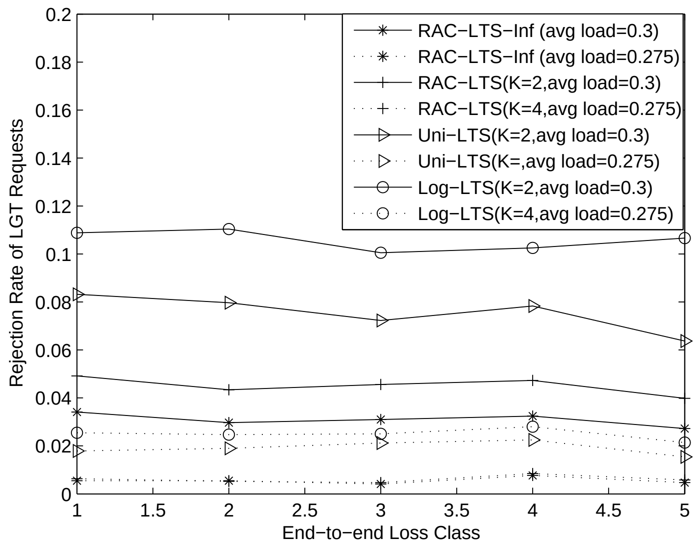

Figure 6.6: Rejection Rate of LGT Requests of Different End-to-End Classes under Different Loss Selection Algorithms

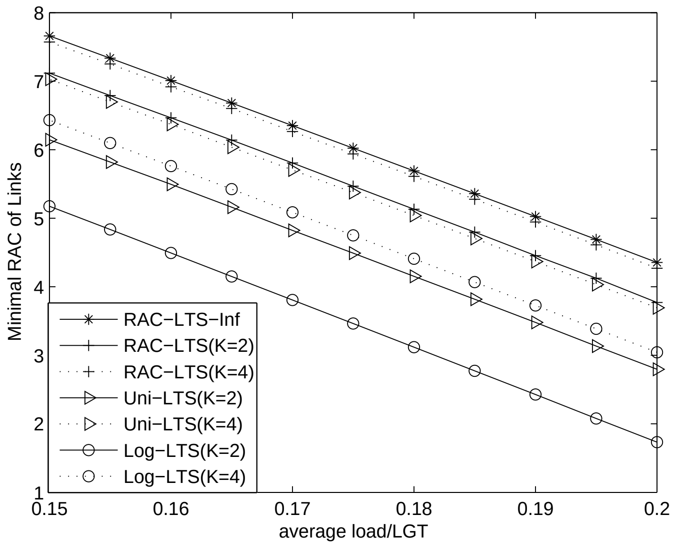

Figure 6.7: Minimal Link RAC in Non-Rejection Scenario under Different Loss Threshold Selection Algorithms

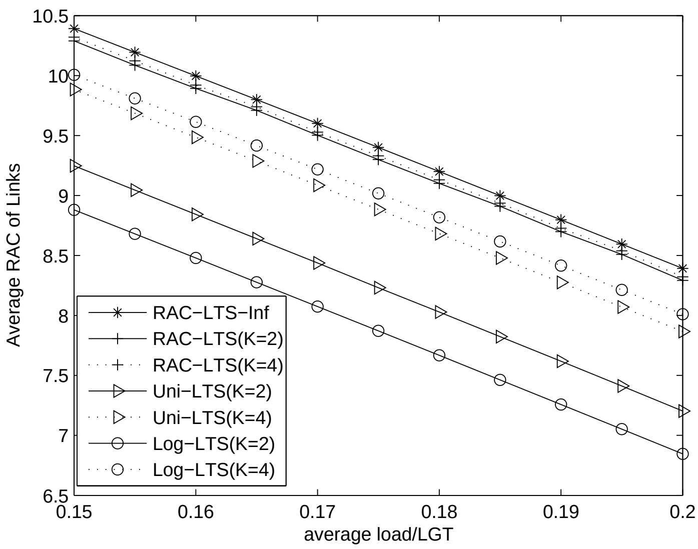

Figure 6.8: Average Link RAC in Non-Rejection Scenario under Different Loss Threshold Selection Algorithms

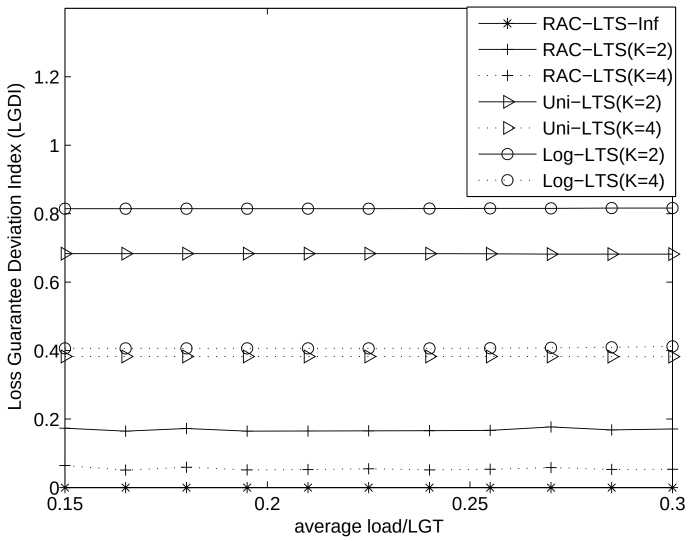

Figure 6.9: Difference Loss Guarantee Deviation Index under Different Loss Threshold Selection Algorithms

6.4 Summary

Based on the concept of RAC, we presented a loss budget partitioning algorithm, RAC-LBP, and a loss threshold selection algorithm, RAC-LTS. Compared with the existing loss budget partitioning algorithms, RAC-LBP is efficient and flexible. RAC-LTS is more efficient than the existing algorithms.

Chapter 7

RAC Based Loss Guaranteed Tunnel Routing Algorithms

In this chapter, we consider the problem of routing loss guaranteed tunnels. To the best of our knowledge, so far there have been no works in the literature on route selection for OBS connections with end-to-end loss guarantees. The research works on providing end-to-end loss guarantees to OBS connections [7][8][11][9][10] assume that the shortest paths are used. The shortest path routing is very likely to create bottleneck links and reduce the network capacity utilization efficiency. In this chapter, we study the problem of routing the LGTs so that the network capacity is utilized efficiently and more LGT requests are admitted. We consider two scenarios of routing, online and offline. In the online scenario, LGT requests arrive one by one and no future demands are known. In the offline scenario, we assume the information of all the LGT requests are given. Such information includes the source-destination node pair and the end-to-end loss requirements of each LGT request. We apply the concept of RAC presented in Chapter 5 in QoS routing in OBS networks. Routing algorithms for the online and offline scenarios are proposed respectively.

7.1 Online Routing Scenario

In this section, we apply the metric of RAC to route selection in OBS networks for the online scenario. In this scenario, we assume that the LGT requests arrive one by one and the future demands are not known.

In the route selection, load balancing can help postpone or avoid the formation of bottleneck links, and thus improve the network resource utilization. A shorter path generally consumes less network resources than a longer path. By taking the load balancing and the preference on shorter paths into consideration, we propose the RAC-based Widest Shortest Path (RAC-WSP) algorithm to route LGT requests in OBS networks. In the RAC-WSP algorithm, if there is only one shortest path between the source and destination, select this shortest path. If there are more than one shortest paths between the source and destination, choose the one with the largest RAC. The RAC of a path is the minimum of the values of RAC of the links along the path. The RAC-WSP algorithm is similar to the traditional WSP algorithm except that the residual capacity is determined using RAC. For the details of the traditional WSP algorithm, please refer to [84]. As RAC estimates the amount of residual capacity more accurately, RAC-WSP algorithm can help improve the network capacity utilization efficiency to accommodate increased number of LGT requests.

7.2 Offline Routing Scenario

In this section, we apply the metric of RAC to route selection in OBS networks for the offline scenario. In this scenario, it is assumed that the source, destination and end-to-end loss requirements of each LGT requests are given, and it is required to route the LGTs in such a way that the number of LGT requests accepted is maximized.

This problem is similar to the multicommodity problem, which has been shown that the only polynomial solution is the linear programming [82]. However, the problem of offline LGT routing we study cannot be solved by the technique of linear programming because the link capacity consumed by several LGTs is not the linear sum of the capacity required by each LGT. Therefore we propose a heuristic algorithm, Residual Admission Capacity based offline Routing algorithm (RAC-OR), as a solution.

RAC-OR works in two phases. Phase I performs initialization. Phase II adopts iterative techniques to improve the route layout. In each iteration, some admitted LGTs are randomly chosen to have their routes re-computed. If the new route layout helps reduce the number of rejected LGT requests, these route layout will be updated. Otherwise, the original route layout is kept. The details of RAC-OR are described below and the pseudo-code is given in the appendix.

7.2.1 RAC-OR Phase I: Initialization

Phase I of RAC-OR performs initialization. We denote the set of all the unadmitted LGT requests as U and the set of all the admitted LGT requests as A. The procedure of Phase I is the following:

1. Let U include all the LGT requests and let A be empty.

2. If U is empty, the algorithm terminates. Otherwise, for each request Q_i in set U, compute the costs of routing Q_i on each of its k shortest paths. The cost indicates the amount of link resources consumed to route the LGT over a specific path. If the path cannot support the end-to-end loss demand of the LGT, the cost is $+\infty$.The procedure to compute the cost of a route is given after the description of RAC-OR Phase II. Among the k shortest paths for request Q_i, the path with the minimal cost is denoted as R_i and the corresponding cost is denoted as C_i.

3. Of all the requests in set U, select the request Q_j whose C_j is the minimal.

   If the minimal C_j is $+\infty$, Phase I terminates and Phase II starts. Otherwise, move the request from set U to set A and route it along the path R_j. The load, weighted loss thresholds and RAC values of the links along path R_j are adjusted accordingly. Then go back to step 2.

7.2.2 RAC-OR Phase II: Iterative Optimization

Phase II of RAC-OR uses iterative techniques to improve the route layout. The detailed procedure is given below:

1. Choose M LGTs randomly and move them from set A to set U. The load, weighted loss thresholds and RAC values of the links traversed by the M LGTs are adjusted accordingly.

2. For each request L_i in set U, compute the costs of routing Q_i on each of its k shortest paths. Among the k shortest paths, the path with the minimal cost is denoted as R_i and the corresponding cost is denoted as C_i.

3. Of all the requests in set U, select the request Q_j whose C_j is the minimal.

   If C_j is not $+\infty$, route the request Q_j along the path R_j and remove request Q_j from set U to set A. The load, weighted loss thresholds and RAC values of the links along path R_j are adjusted accordingly. Then Then go back to step 2.

   If C_j is $+\infty$, compute the number of LGT requests in set U. If the number increases comparing to last iteration, restore the route layout to that at the end of last iteration. Otherwise keep the changes made in this iteration. Here ends this iteration. Then we check whether the stopping criteria are met. If the stopping criteria are not met, the algorithm goes back to step 1 to start another iteration. If one of the following conditions are met, the algorithm finishes:

   - Set U is empty.
   - the number of rejected LGT requests has not reduced for a given number of iterations.
   - the number of iterations has exceeded a pre-set threshold.

When RAC-OR finishes, set A is the set of all the admitted LGT requests.

There are two principal intuitions for RAC-OR. First, the LGT requests consuming less network resources are routed first to maximize the number of LGT requests accommodated. Secondly, when routing an LGT, select the route from the k shortest paths. We confine the possible route candidates of an LGT request to the k shortest paths because of the difficulty to determine the cost of using a link before the route is known. In IP networks, the QoS requirements of a connection are usually translated into the demand of bandwidth whose consumption is the same across all the links. So the cost of each link is known before the route is determined. However, in OBS networks, for an LGT request, before the route is decided, the values of loss guarantees and the amount of capacity consumption on the links cannot be decided. As shorter routes usually consume less network resources, it is very likely that the most efficient route for an LGT request is one of the k shortest paths. Therefore, for every LGT request, RAC-OR considers the k shortest paths only.

7.2.3 Cost of Routing an LGT

Now we explain how RAC-OR measures the cost of routing an LGT over a specific route. The following expression describes the link capacity consumption conditions in network net.

$$Res(net) = \sum_{l \in net} \frac{1}{\beta_l}$$

Here β_l is the RAC value on link l. The smaller β_l is, the bigger $\frac{1}{\beta_l}$ is. The value of $\frac{1}{\beta_l}$ increases dramatically when the capacity in link l is close to be used up. Therefore, a smaller value of $Res(net)$ indicates more link residual capacity and a more balanced link load distribution in the network.

We define the cost of routing an LGT request over a path as the increment of $Res(net)$ it causes. In details, the cost of routing LGT request Q over the path R, $Cost(Q, R)$. is computed as follows. First, use the loss budget partitioning algorithm in [10] to compute the loss guarantee value on each link l, T_l. This algorithm also determines whether the path can support the end-to-end loss demand of the LGT request. Then:

1. If path R cannot support the end-to-end loss requirements of LGT request Q, $Cost(Q, R)$ $=+\infty$
2. If path R has sufficient capacities to satisfy the end-to-end loss requirements of LGT request Q, compute the RAC of each link along the path before and after the request Q is admitted with the loss threshold of T_l. $Cost(Q, R) = \sum\limits_{l \in R} (\frac{1}{\beta_l^{old}} - \frac{1}{\beta_l^{new}})$. Here β_l^{new} and β_l^{old} are the RAC value on link l after and before LGT request Q is routed on the path R, respectively.

7.3 Numerical Results

This section evaluates the effectiveness of the proposed routing algorithms in terms of the rejection rate of the LGT requests. We suppose that the network under investigation provides 5 end-to-end loss guarantee classes, with the end-to-end loss guarantees equal to 1×10^{-4}, 5×10^{-4}, 1×10^{-3}, 5×10^{-3} and 1×10^{-2}, respectively. Between each node pair, there are 5 LGT requests, corresponding to the 5 end-to-end loss guarantee classes. In the online routing scenario, the LGT requests are fed into the network one by one in a random sequence. After the route is selected, the loss budget partitioning algorithm in [10] is used to determine whether the path has enough residual capacity to admit this LGT request. Once an LGT request is accepted, it will stay there.

In our study, we consider the 14-node NSFNET network and the 33-node Pan-European network (see http://www.geant.net). In the Pan-European network, some links are added to make the network bi-connected. In the NSFNET network, each link carries 16 wavelengths, while in the Pan-European network, each link carries 64 wavelengths. We study the performance of the routing algorithms by varying the average maximum permissible load (which will be simply referred as "average load" in the rest of the chapter) of the LGT requests. For a given average load c, the load of each LGT request is uniformly distributed between $1.5c$ Erlang and $0.5c$ Erlang. For each average load, 20 sets of LGT requests are generated randomly. On each link, the loss thresholds, namely the possible values of loss guarantees, are 2×10^{-5}, 8×10^{-5}, 3.2×10^{-4}, 1.28×10^{-3} and $5.12\times10^{-3,}$ respectively.

We compare the performance of the following algorithms.

1. Shortest path algorithm (referred as SPF), where the minimum-hop path is selected. If there are more than one shortest path existing between the source and the destination nodes, choose one randomly.

2. Widest shortest path algorithm where the link residual capacity is determined in the method of computing residual bandwidth in IP networks (referred as EC-WSP).

3. Widest shortest path algorithm where the link residual capacity is determined by the effective load over the link (referred as load-WSP). In this algorithm, a link carrying less effective load is considered to have larger residual capacity .

4. RAC-WSP proposed in this chapter.

5. Ordered RAC-WSP (referred as O-RAC-WSP). Here LGT requests are routed using the RAC-WSP algorithm and they are considered in a increasing order of their capacity stringency, η, which is defined as follows:

$$\eta = \rho \div \gamma \times \varphi \tag{7.1}$$

where φ is the hop number of the shortest path, ρ and γ are the load and the required end-to-end loss guarantee, respectively. The rationale is to route the LGTs with lower capacity requirements first to increase the total number of LGTs accepted finally. We found that such an order achieves the best performance compared with the other routing orders we have tested, including the random order, the least-load-first order, the largest-load-first order, the longest-flow-first order, the shortest-flow-first order and the most-stringent-first order.

6. RAC-OR proposed in this chapter. The parameters in RAC-OR are set as follows:

   - The number of the shortest paths considered to compute the routing cost for each LGT request, k, is 5. Different values of k in both network topologies have been tested. A larger value of k gives a better performance. However, when k is greater than 5, the performance improvement as the k increases very slowly while the computation intensity continues to grow. So we set k as 5.
   - In Phase II, in each iteration, 10% of the LGTs in set A requests are moved from set A to set U .
   - The iterative optimization in Phase II stops when the number of iterations has reached 100, or the performance has not enhanced for 10 iterations, or the set U is empty.

The algorithms of SPF, EC-WSP, load-WSP and RAC-WSP work in the online scenario, while the algorithms of O-RAC-WSP and RAC-OR work in the offline scenario. Figure 7.1 shows the average rejection rates of LGT requests in the NSFNET network. Figures 7.2 and 7.3 show the rejection rates in each of the 20 trials, when the average load is low and high, respectively, also in the NSFNET network. Figures 7.4, 7.5 and 7.6 show the corresponding results in the Pan-European network.

We can observe that, in the online routing scenario, RAC-WSP algorithm rejects less LGT requests than the other three online routing algorithms. The advantage of RAC-WSP is greater when the network load is heavier. The performances of load-WSP and EC-WSP are almost the same as SPF. It shows that it is very inaccurate to use the traditional methods such as residual bandwidth or load to determine the amount of link residual capacity remaining.

In the offline scenario, RAC-OR gives a better performance than O-RAC-WSP. It shows that RAC-OR can utilize the network resources more effectively when deploying the LGTs. The advantage of RAC-OR over the other algorithms is greater when the network load is heavier. In practice, the network operator usually has the information of the LGT requests, as users demanding end-to-end loss guarantees usually sign SLA with the operator. So RAC-OR can help the network operator to utilize the network capacity effectively and maximize the revenue. The rejection rates of 20 trials presented in Fig 7.2, 7.3, 7.5 and 7.6 show that our algorithms give statistically different (better) performance comparing to the other algorithms.

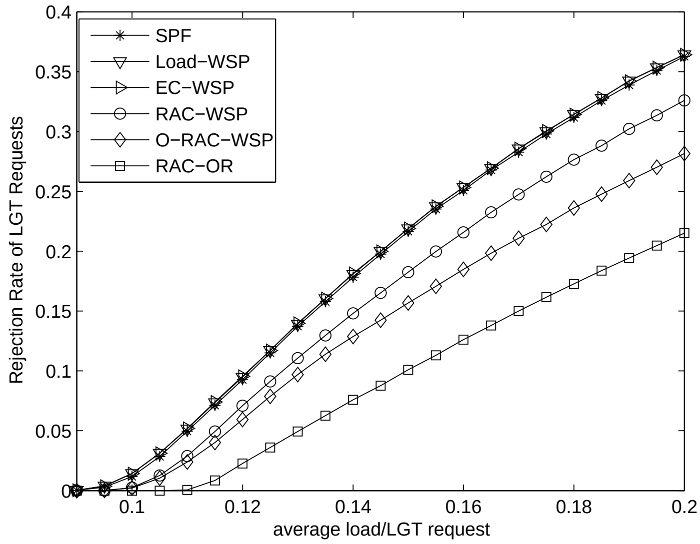

Figure 7.1: Rejection Rates by Different Routing Algorithms (NSFNET Network)

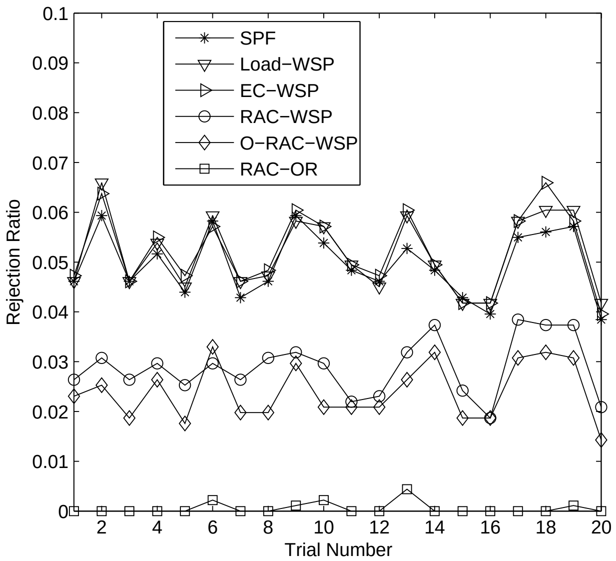

Figure 7.2: Rejection Rates by Different Routing Algorithms (NSFNET Network, Average Load=0.11 Erlang)

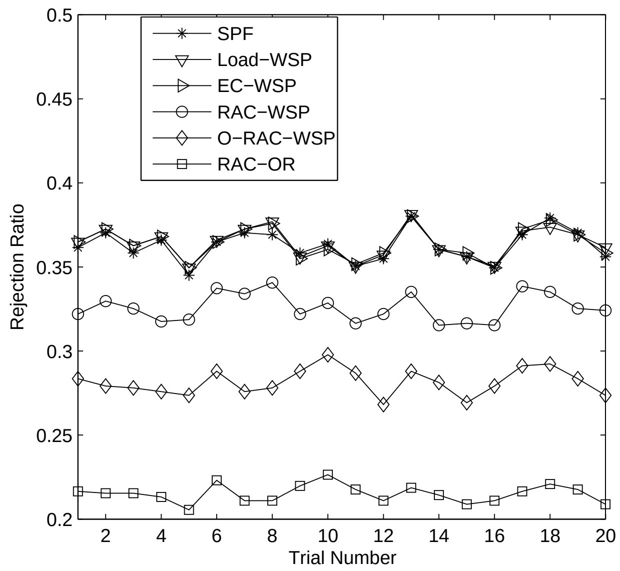

Figure 7.3: Rejection rates by Different Routing Algorithms (NSFNET Network, Average Load=0.2 Erlang)

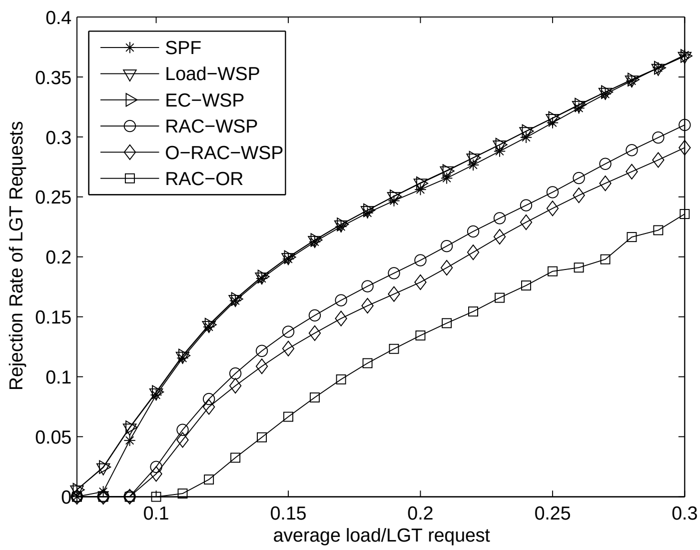

Figure 7.4: Rejection Rates by Different Routing Algorithms (Pan-European Network)

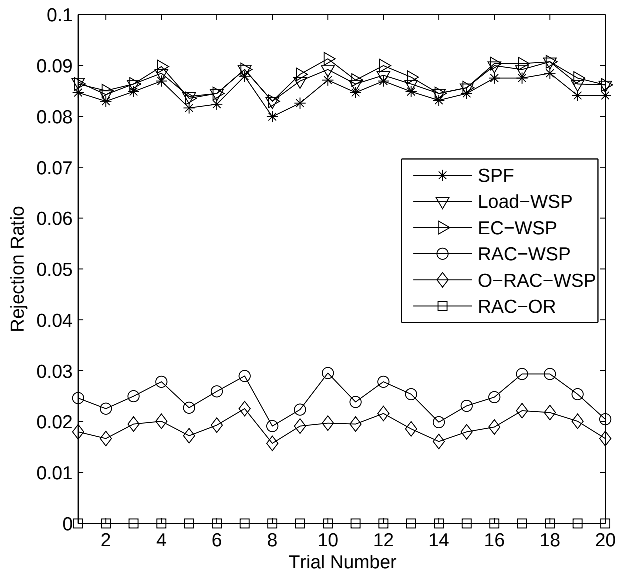

Figure 7.5: Rejection Rates by Different Routing Algorithms (Pan-European Network, Average Load=0.1 Erlang)

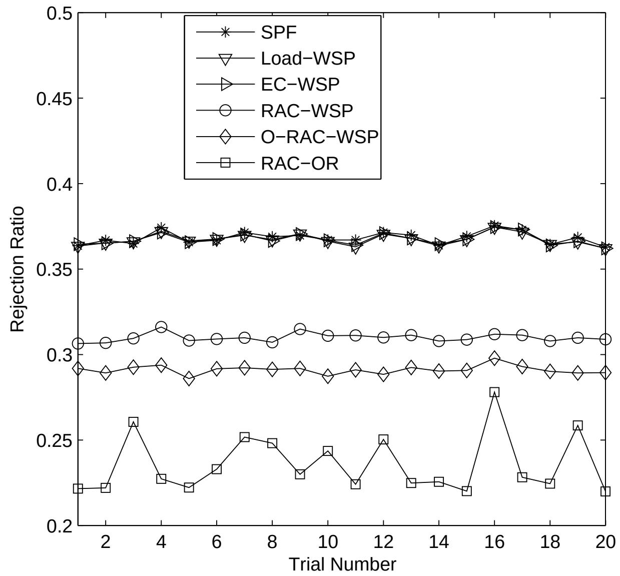

Figure 7.6: Rejection Rates by Different Routing Algorithms (Pan-European Network, Average Load=0.3 Erlang)

7.4 Summary

We considered the problem of LGT routing in OBS networks. Based on the concept of RAC, we presented the routing algorithm for the online scenario and the offline scenario. For the online scenario, we presented a routing algorithm, RAC-WSP, which is a widest shortest path algorithm that uses RAC as a measure of residual capacity. For the offline scenario, we presented an algorithm, RAC-OR. Experiment results showed that our algorithms can admit more LGTs compared to other algorithms.

Chapter 8

Conclusions and Future Work

8.1 Research Contribution

The ever-increasing Internet traffic results in a need for an efficient data transmission paradigm over optical networks. OBS emerges as an effective and practical transmission paradigm over WDM networks and has attracted a lot of research interests. In this thesis, we analyzed two unique features of the OBS networks which arise from the bufferless nature of the core nodes. The first feature is the streamline effect. We showed that the traditional Erlang B formula is inaccurate due to the ignorance of this effect. We also presented a new loss estimation formula with the effect taken into consideration. The second feature is the way of measuring link residual capacity. We showed that the traditional methods of computing residual bandwidth in IP or ATM networks is not applicable to OBS networks due to its bufferless nature. We introduced a new metric, called residual admission capacity (RAC) and proposed a method to compute it. We showed that our metric gives a more accurate estimation on the amount of link residual capacity.

We applied the loss estimation formula presented in the offline route layout optimization for best-effort traffic. We first considered the route layout optimization to minimize the overall burst loss. We then considered the route layout design for fast and efficient failure recovery. Two link-disjoint routes, primary and backup path, are determined for each flow, with the objective to minimize the expected burst loss. MILP formulations were developed for the above problems. Since it needs intensive computation to solve MILP problems, heuristic algorithms were proposed. Numerical results show that the proposed algorithms can find route layouts with lower burst loss rates than the algorithms where the loss estimation is based on the Erlang B formula.

We applied the metric of RAC in connection configurations for loss-guaranteed traffic. Based on the concept of RAC, we designed an RAC based loss budget partitioning algorithm (RAC-LTS). It aims to partition the loss budget in such a manner that the minimal RAC along the path is maximized to defer the formation of bottleneck links. It effectively utilizes the link capacity and can work with different loss threshold settings. We also proposed an RAC based loss threshold selection algorithm (RAC-LTS). This algorithm solves the loss threshold selection problem in two phases. In Phase I, the constraint on the number of loss threshold values is temporarily removed and loss thresholds can take any positive real value. We determine the loss guarantees for each LGT in each traversing link in such a way that the minimal RAC along the path after the LGT being deployed is maximized. Phase II is the process of loss threshold quantization. In Phase II, we determine the set of loss thresholds which minimize the extra resource consumption due to the constraint on the finity of the loss threshold number. Experiment results showed that RAC-LTS is more effective than the other loss threshold selection algorithms. With the loss thresholds selected by RAC-LTS, more LGT requests can be admitted.

The metric of RAC was also applied in the traffic engineering for loss-guaranteed traffic. Two routing algorithms, RAC-WSP and RAC-OR, were proposed for the online and offline scenarios respectively. The routing algorithms use RAC as the metric to evaluate the link residual capacity and can make better decision in route selection. The experiment results show that the algorithms presented by us can admit more LGT requests than the other algorithms.

8.2 Future Work

1. The research work in this thesis assumes single-path routing, where a flow or an LGT takes only one LSP to forward the bursts. However, it is shown that multi-path routing can achieve better performance in terms of burst loss as it provides more flexibility in route selection. Therefore, the work in this thesis can be extended to multi-path routing in the future. Multi-path routing has the drawback of possible out-of-sequence bursts, which needs to be addressed by any solutions.

2. In this thesis, the concept of RAC is used to design algorithms to select loss thresholds and route LGTs. However, we consider these two problems separately. In studying the problem of loss threshold selection, we assumed the routes are fixed. While in studying the route problems, we assumed that the loss thresholds are given. However, joint selection of routes

and loss thresholds are expected to result in a better performance. It is a problem open for future work.

3. This thesis considers the route selection problem for the best-effort and loss-guaranteed traffic respectively. However, in practice, these two types of traffic can coexist within a network. The research work in this thesis can be extended to design traffic engineering methods in such a multi-service OBS network. The objective could be to maximize the revenue or maximize the network efficacy function. The concept of network efficacy function was proposed in [85], where a utility function is used to describe a network user's level of satisfaction to the service provided and the network efficacy function is the summation of the utility functions of all the users. Research on this problem is important to provide multiple services in OBS networks.

Appendix A:

Explanation on the Approximate Relationship $\sum_{m=1}^{D} g_m \leq \gamma$

The exact relationship between the loss guarantee on the mth link, g_m, and the end-to-end loss requirements, γ, should be that $1 - \prod_{m=1}^{D}(1 - g_m) \leq \gamma$. We have the following equation by expanding $\prod_{m=1}^{D}(1 - g_m)$,

$$\prod_{m=1}^{D}(1 - g_m) = 1 - \sum_{m=1}^{D} g_m + \sum_{m,n=1,m\neq n}^{D} g_m g_n - \sum_{m,n,o=1,m\neq n\neq o}^{D} g_m g_n g_o + ... \qquad (1)$$

Since those loss-guaranteed services require relatively low end-to-end loss rate, γ always takes a small positive value, so g_m are also small positive, very close to 0. So we can ignore the second and higher-order terms in right side of Eq(1). As a result,

$$\prod_{m=1}^{D}(1 - g_m) \approx 1 - \sum_{m=1}^{D} g_m$$

$$1 - \prod_{m=1}^{D}(1 - g_m) \approx \sum_{m=1}^{D} g_m$$

So the exact relationship $1 - \prod_{m=1}^{D}(1 - g_m) \leq \gamma$ is approximately equal to $\sum_{m=1}^{D} g_m \leq \gamma$. This approximation turns a complicated multiplication formulation into a linear summation one and makes the computation simpler. So we use this approximate relationship in this thesis.

Appendix B:

Pseudo code of RAC-OR

U =set of all LGT requests;

$A = \phi$;

$STOP\_OR\_NOT$ =FALSE;

$Iteration = 0$;

$Prev\_Rej\_No$ =Number of all the LGT requests;

$T1$ = maximal number of iterations;

$T2$ = maximal number of iterations without performance enhancements;

while (STOP_OR_NOT==FALSE)

{

$min\_c = +\infty$; $min\_req = -1$;

for (each LGT request Q_i in set U)

{

```
    C_i = +∞;
    for ( each of Q_i's K shortest path, R_{i,k})
    {
        C_{i,k}=cost (Q_i , R_{i,k});
        if (C_{i,k} < C_i)
            {   C_i = C_{i,k};        R_i = R_{i,k};   }
        else;
    }
    if (C_i < min_c)
        {   min_c = C_i;    min_req = i;       }
    else;
}
if (min_c == +∞)
{
```

Rej_No =the number of LGT requests in set U;

if ($Rej\_No < Prev\_Rej\_No$)

{

$Enhanced\_Iteration = Iteration$;

$Prev\_Rej\_No = Rej\_No$;

}

else restore the route layout to that in the end of last iteration;

$Iteration = Iteration + 1$;

if (Rej_No==0 || Iteration > T1 || Iteration - Enhanced_Iteration > T2)

STOP_OR_NOT==TRUE;

else

M LGTs are randomly chosen to be removed from set A to set U;

}

else

{ route the LGT request Numbered min_req on route R_{min_req} and move the request from set U to set A; }

}

Publication

International Journal Papers

1. Q. Chen, G. Mohan, and K. C. Chua, "Route Optimization in Optical Burst Switched Networks Considering Streamline Effect," Elsevier Computer Networks Journal, Vol. 52, No. 10, pp. 2033-2044, July 2008.

2. Q. Chen, G. Mohan, and K. C. Chua, "Residual Admission Capacity in Optical Burst Switching Networks and its Application in Configuration of Loss Guaranteed Tunnels," to appear in IEEE/OSA Lightwave Technology Journal.

3. Q. Chen, G. Mohan, and K. C. Chua, "Residual Admission Capacity in Optical Burst Switching Networks and its Application in QoS Routing," to be submitted.

International Conference Papers

1. Q. Chen, G. Mohan, and K. C. Chua, "Offline Route Optimization Considering Streamline Effect in OBS Networks," in Proceedings of IEEE ICC-2006, June 2006.

2. Q. Chen, G. Mohan, and K. C. Chua, "Route Optimization for Efficient Failure Recovery in Optical Burst Switched Networks," in Proceedings of IEEE HPSR-2006, June 2006.

3. Q. Chen, G. Mohan, and K. C. Chua, "Residual Admission Capacity in Optical Burst Switching Networks and its Application in Routing Loss-Guaranteed Flows," in Proceedings of IEEE LCN-2006, November 2006.

Bibliography

[1] C. Qiao and M. Yoo, "Optical Burst Switching - A New Paradigm for an Optical Internet", Journal of High Speed Networks, Special Issue on Optical Networks, Vol. 8, No. 1, pp. 69-84, 1999.

[2] J. S. Turner, "Terabit burst switching", Journal of High Speed Networks, Vol. 8, pp. 3-16, 1999.

[3] Y. Xiong, M. Vandenhoute and H. C. Cankaya, "Control Architecture in Optical Burst-Switched WDM Networks", IEEE Journal on Selected Areas in Communications, Vol. 18, No. 10, pp. 1838 - 1851, 2000.

[4] M. Yoo, C. Qiao and S. Dixit, "Optical Burst Switching for Service Differentiation in the Next Generation Optical Internet", IEEE Communications Magazine, Vol. 39, No. 2, pp. 98-104, 2001.

[5] K. Thompson, G. J. Miller and R. Wilder, "Wide-Area Internet Traffic Patterns and Characteristics", IEEE Network, Vol. 11, No. 6, pp.10 - 23, 1997.

[6] J. Teng and G. N. Rouskas, "A Traffic Engineering Approach to Path Selection in OBS Networks,." OSA Journal of Optical Networking, Vol. 4, No. 11, pp. 759-777, 2005.

[7] S. Kim, K. Min and M. Kang, "Dynamic Wavelength Management Method in OBS Networks", U.S. Patent Application 10/338 243, 2003.

[8] Q. Zhang, V. Vokkarane, J. P. Jue and B. Chen, "Absolute QoS Differentiation in Optical Burst-Switched Networks", IEEE Journal on Selected Areas in Communications, Vol. 22, No. 9, pp. 1781-1795, 2004.

[9] M.H. Phung, K. C. Chua, G. Mohan, M. Motani and T.C. Wong, "Absolute QoS Signalling and Rreservation in Optical Burst-Switched Networks", in Proceedings of Global Communications Conference, 2004.

[10] M.H. Phung, K. C. Chua, G. Mohan, M. Motani and T.C. Wong, "An Absolute QoS Framework for Loss Guarantee in Optical Burst Switched Networks", IEEE Transaction on Communications, Vol. 55, No. 6, pp. 1191-1201, 2007.

[11] L. Yang and G. N. Rouskas, "A Framework for Absolute QoS Guarantees in Optical Burst Switched Networks", in Proceedings of International Conference on Broadband Networks, 2006.

[12] R. Braden et al., "Integrated Services in the Internet Architecture: an Overview", RFC 1633, 1994.

[13] S. Blake et al., "An Architecture for Differentiated Services", RFC 2475, 1998.

[14] I. Widjaja, "Performance Analysis of Burst Admission-Control Protocols", in Proceedings of IEE Communications, 1995.

[15] E. Varvarigos and V. Sharma, "The ready-to-go virtual circuit protocol : A Loss-Free Protocol for Multigigabit Networks using FIFO Buffers", IEEE/ACM Transactions on Networking, Vol. 5, No. 5, pp. 705-718, 1997.

[16] J. Y. Wei and R. I. McFarland, "Just-in-time signaling for WDM Optical Burst Switching Networks", IEEE/OSA Lightwave Technology Journal, Vol. 18, No. 12, pp. 2019-2037, 2000.

[17] J. Y. Wei, J. L. Pastor, R. S. Ramamurthy and Y. Tsai, "Just-in-Time Optical Burst Switching for Multiwavelength Networks", in Proceedings of IFIP Broadband Communications, 1999.

[18] M. Yoo and C. Qiao, "Just-Enough-Time(JET): a High Speed Protocol for Bursty Traffic in Optical Networks", Digest of IEEE/LEOS Summer Topical Meetings on Technologies for a Global Information Infrastructure, 1997.

[19] S. Sheeshia and C. Qiao, "Synchronous Optical Burst Switching", in Proceedings of International Conference on Broadband Communications, Networks and Systems, 2004.

[20] J. Ramamirtham and J. Turner, "Time Sliced Optical Burst Switching", in Proceedings of IEEE Conference on Computer Communications, 2003.

[21] Z. Zhang, L. Liu and Y. Yang, "Slotted Optical Burst Switching (SOBS) Networks", in Proceedings of IEEE International Symposium on Network Computing and Applications, 2006

[22] Y. Liu, G. Mohan and K. C. Chua , "A Dynamic Bandwidth Reservation Scheme for a Collision-Free Time-Slotted OBS Network," in Proceedings of IEEE/CreateNet Workshop on Optical Burst/Packet Switching, 2005

[23] A. Rugsachart and R. A. Thompson, "An Snalysis of Time-Synchronized Optical Nurst Switching", in Proceedings of High Performance Switching and Routing, 2006

[24] L. Ou, X. Tan, Y. Ma and Z. Yang, "A Framework to Evaluate Blocking Performance of Time-Slotted Optical Burst Switched Networks", in Proceedings of Local Computer Networks, 2005

[25] J. Xu, C. Qiao, J. Li and G. Xu, "Efficient Burst Scheduling Algorithms in Optical Burst-Switched Networks using geometric techniques", IEEE Journal on Selected Areas in Communications, Vol. 22, No. 9, pp. 1796 - 1811, 2004

[26] S. K. Tan, G. Mohan and K. C. Chua, "Algorithms for Burst Rescheduling in WDM Optical Burst Switching Networks", Journal of Computer Networks, Vol. 41, No. 1, pp. 41-55, 2003.

[27] A. Detti, V. Eramo and M. Listanti, "Optical Burst Switching with Burst Drop (OBS/BD): an Easy OBS Improvement", in Proceedings of International Conference on Communications, 2002

[28] A. Detti, V. Eramo and M. Listanti, "Performance Evaluation of a New Technique for IP Support in a WDM Optical Network: Optical Composite Burst Switching (OCBS)", IEEE/OSA Lightwave Technology Journal, Vol. 20, No. 2, pp. 154 – 165, 2002

[29] V. M. Vokkarane and J. P. Jue, "Prioritized Burst Segmentation and Composite Burst Assembly Techniques for QoS Support in Optical Burst-Switched Networks", IEEE Journal on Selected Areas in Communications - Special Issue on High-Performance Optical/Electronic Switches/Routers for High-Speed Internet, Vol. 21, No. 7, pp. 1198-1209, 2003

[30] V. M. Vokkarane and J. P. Jue, "Segmentation-Based Nonpreemptive Channel Scheduling Algorithms for Optical Burst-Switched Networks", IEEE/OSA Lightwave Technology Journal, Vol. 23, No. 10, pp. 3125 – 3137, 2005

[31] J. Li and C. Qiao,"Schedule Burst Proactively for Optical Burst Switching Networks", in Proceedings of Global Communications Conference, 2003

[32] N. Barakat and E. H. Sargent, "Separating Resource Reservations from Service Requests to Improve the Performance of Optical Burst-Switching Networks", IEEE Journal on Selected Areas in Communications, Vol. 24, No. 4, pp. 107, 2006

[33] C. Y. Li, G. M. Li, P.K.A Wai and V.O.K. Li, "Novel Resource Reservation Schemes for Optical Burst Switching", in Proceedings of IEEE International Conference on Communications, 2005

[34] M. H. Phung, K. C. Chua, G. Mohan, M. Motani, T. C. Wong and P. Y. Kong, "On Ordered Scheduling for Optical Burst Switching", Journal of Computer Networks, vol. 48, no. 6, pp 891-909, August 2005.

[35] J. Li, C. Qiao, J. Xu and D. Xu, "Maximizing Throughput for Optical Burst Switching Networks", IEEE/ACM Transactions on Networking, Vol. 15, No. 5, pp. 1163 - 1176, 2007

[36] J. Li and C. Qiao, "Recent Progress in the Scheduling Algorithms in Optical-Burst-Switched Networks", Journal of Optical Networking, Vol. 3, No.4, pp. 229-241, 2004

[37] J. Li, G. Mohan and K. C. Chua, "Dynamic Load Balancing in IP-over-WDM Optical Burst Switching Networks," Journal of Computer Networks, Vol. 47, No.3, pp. 393-408, 2005.

[38] J. Lu, G. Mohan, K.C. Chua and Y. Liu, "Optimal Multi-Path Traffic Engineering in Optical Burst Switching Networks", in Proceedings of IEEE Workshop on High Performance Switching and Routing, 2006.

[39] S. Ganguly, S. Bhatnagar, R. Izmailov and C. Qiao, "Multi-Path Adaptive Optical Burst Forwarding", in Proceedings of High Performance Switching and Routing, 2004

[40] X. Wang, H. Morikawa and T. Aoyama, "Burst Optical Deflection Routing Protocol for Wavelength Routing WDM Networks", in Proceedings of SPIE/IEEE Optical Networking and Communication Conference, 2000.

[41] S. Kim, N. Kim and M. Kang, " Contention Resolution for Optical Burst Switching Networks Using Alternative Routing", in Proceedings of IEEE International Conference on Communications, 2002

[42] S. Lee, K. Sriram, H. Kim and J. Song, "Contention-Based Limited Deflection Routing Protocol in Optical Burst-Switched Networks", IEEE Journal on Selected Areas in Communications, Vol. 23, No. 8, pp. 1596 – 1611, 2005

[43] M. Jin, O. Yang, "On controlling the loss rate in optical burst switching networks", in Proceedings of International Workshop on Optical Burst/Packet Switching, 2006

[44] G. Mohan, K. Akash and M. Ashish, "Efficient Techniques for Improved QoS Performance in WDM Optical Burst Switched Networks," Computer Communications Journal, vol. 28, no. 7, pp 754-764, 2005.

[45] D.Q. Liu and M. T. Liu, "Differentiated Services and Scheduling Scheme in Optical Burst-Switched WDM Networks", in Proceedings of IEEE International Conference on Communications, 2002

[46] M. Yoo, C. Qiao and S. Dixit, "QoS Performance of Optical Burst Switching in IP-over-WDM Networks", IEEE Journal on Selected Areas in Communications,Vol. 18, No. 10, pp. 2062 – 2071, 2000

[47] A. Kaheel and H. Alnuweiri, "Strict Priority Scheme for Quality-of-Service Provisioning in Optical Burst Switching Networks", in Proceedings of IEEE Symposium on Computers and Communications, 2003.

[48] W. Liao and C .H. Loi, "Providing Service Differentiation for Optical-Burst-Switched Networks", IEEE/OSA Lightwave Technology Journal, Vol. 22, No. 7, pp. 1651 - 1660, 2004

[49] L. Yang, Y. Jiang and S. Jiang, "A Probabilistic Preemptive Scheme for Providing Service Differentiation in OBS Networks", in Proceedings of Global Communications Conference, 2003.

[50] Y. Chen, C.g Qiao, M Hamdi and D.H.K. Tsang,"Proportional Differentiation: A Scalable QoS Approach", IEEE Communications Magazine, Vol. 41, No. 6,pp. 52 - 58, 2003

[51] H.C. Cankaya, S. Charcranoon and T.S. El-Bawab, "A Preemptive Scheduling Technique for OBS Networks with Service Differentiation", in Proceedings of Global Communications Conference, 2003

[52] S. K. Tan, G. Mohan and K. C. Chua, "Feedback-based Offset Time Selection for End-to-End Proportional QoS Provisioning in WDM Optical Burst Switching Networks," Journal of Computer Communications, vol. 30, no. 4, pp. 904-921, 2007.

[53] J. Phuritatkul, Y. Ji and S. Yamada, "Proactive Wavelength Pre-Emption for Supporting Absolute QoS in Optical-Burst-Switched Networks", IEEE/OSA Lightwave Technology Journal, Vol. 25, No. 5, pp. 1130 – 1137, 2007

[54] C. W. Tan, G. Mohan and J.C.S Lui, "Achieving Multi-Class Service Differentiation in WDM Optical Burst Switching Networks: A Probabilistic Preemptive Burst Segmentation Scheme", IEEE Journal on Selected Areas in Communications, 2006,Vol. 24, No. 12, pp. 106 - 119, 2006

[55] A. Maach, G. Bochmann and H. Mouftah, "Robust Optical Burst Switching", in Proceedings of Telecommunications Network Strategy and Planning Symposium, 2004

[56] M. H. Phung, D. Shan, K. C. Chua and G. Mohan, "Performance Analysis of a Bufferless OBS Node Considering the Streamline Effect,"IEEE Communications Letters, vol. 10, no. 4, pp. 293-295, April 2006.

[57] K. Kar, M. Kodialam and T.V. Lakshman, "Minimum Interference Routing of Bandwidth Guaranteed Tunnels with MPLS Route Layout Applications", IEEE Journal on Selected Areas in Communications, Vol. 18, No. 12, pp. 2566 - 2579, 2000

[58] A. Elwalid, C. Jin, S. Low and I. Widjaja, "MATE: MPLS Adaptive Traffic Engineering", in Proceedings of IEEE Conference on Computer Communications, 2001.

[59] E. Kozlovski, "Survivability of Wavelength-Routed Optical Burst-Switched Networks with Guaranteed IP Services", in Proceedings of 5th International Conference on Transparent Optical Networks, 2003

[60] Yufeng Xin et al, "Fault Management with Fast Restoration for Optical Networks", in Proceedings of International Conference on Broadband Communications, Networks and Systems, 2004

[61] Yufeng Xin et al, "A Novel Fast Restoration Mechanism for Optical Burst Switched Network", in Proceedings of Worksop on Optical Switching, 2004

[62] D. Griffith and S. K. Lee, "A 1+1 Protection Architecture for Optical Burst Switched Networks", IEEE Journal on Selected Areas in Communications, Vol. 21, No. 9, pp. 1384-1398, 2003

[63] A. Elwalid, D. Mitra, I. Sanieec and I. Widjaja, “Routing and Protection in GMPLS Networks: from Shortest Paths to Optimized Designs”, IEEE/OSA Lightwave Technology Journal, Vol 21, No. 11, pp. 2828 - 2838, 2003

[64] S. Ramamurthy, L. Sahasrabuddhe and B. Mukherjee, ”Survivable WDM Mesh Networks”, IEEE/OSA Lightwave Technology Journal, Vol. 21, No. 4, pp. 870 – 883, 2003

[65] H. Zang, C. Ou and B. Mukherjee, “Path-Protection Routing and Wavelength Assignment (RWA) in WDM Mesh Networks under Duct-Layer Constraints”, IEEE/ACM Transactions on Networking, Vol. 11, No. 2, pp. 248 – 258, 2003

[66] L. Sahasrabuddhe, S. Ramamurthy and B. Mukherjee, “Fault Management in IP-over-WDM Networks: WDM Nrotection versus IP Restoration”, IEEE Journal on Selected Areas in Communications, Vol. 20, No. 1, pp. 21 - 33, 2002

[67] D.A. Dunn, W.D. Grover and M.H. MacGregor, “Comparison of K-Shortest Paths and Maximum Flow Routing for Network Facility Restoration”, IEEE Journal on Selected Areas in Communications, Vol. 12, No. 1, pp. 88 - 99, 1994

[68] R. Guerin, H. Ahmadi and M. Naghshineh, “Equivalent Capacity and its Application to Bandwidth Allocation in High-Speed Networks”, IEEE Journal on Selected Areas in Communications, Vol. 9, No. 7, pp. 968 - 981, 1991.

[69] X. Yu, Y. Chen and C. Qiao, “Study of Traffic Statistics of Assembled Burst Traffic in Optical Burst Switched Networks,”, in Proceedings of Optical Networking and Communications Conference, 2002.

[70] Z. Rosberg, L. V. Hai, M. Zukerman and J. White, "Performance Analyses of Optical Burst-Switching Networks", IEEE Journal on Selected Areas in Communications, Vol. 21, No. 7, pp. 1187 - 1197, 2003

[71] M.C.F. De Toledo and W. Zucchi, "Simulation of an Optical Burst Switch Using Fiber Delay Lines", in Proceedings of SBMO/IEEE MTT-S International Conference on Microwave and Optoelectronics, 2005

[72] C. F. Hsu, T. L. Liu and N. F. Huang, "On the Deflection Routing in QoS Supported Optical Burst-Switched Networks", in Proceedings of IEEE International Conference on Communications, 2002

[73] H. M. H. Shalaby, "A Simplified Performance Analysis of Optical Burst-Switched Networks", IEEE/OSA Lightwave Technology Journal, Vol. 25, No. 4, pp. 986 - 995, 2007

[74] M.T. Anan, G.M. Chaudhry and D. Benhaddou, "Architecture and Performance of A Next-Generation Optical Burst Switch (OBS)", in Proceedings of International Conference on Broadband Communications, Networks and Systems, 2006

[75] B. Kantarci, S. Oktug and T. Atmaca, "Analyzing the Effects of Burst Assembly in Optical Burst Switching under Self-Similar Traffic", in Proceedings of Advanced International Conference on Telecommunications, 2005

[76] J.Phuritatkul, Y. Ji and Y. Zhang, "Blocking Probability of a Preemption-Based Bandwidth-Allocation Sscheme for Service Differentiation in OBS Networks", IEEE/OSA Lightwave Technology Journal, Vol. 24, No. 8, pp. 2986 - 2993, 2006

[77] X. Yu, J. Li, X. Cao, Y. Chen and C. Qiao, "Traffic Statistics and Performance Evaluation in Optical Burst Switched networks ", IEEE/OSA Lightwave Technology Journal, Vol. 22, No. 12, pp. 2722 – 2738, 2004

[78] S. Azodolmolky, A. Tzanakaki and I. Tomkos, "Study of the Impact of Burst Assembly Algorithms in Optical Burst Switched Networks with Self-Similar Input Traffic", in Proceedings of International Conference on Transparent Optical Networks, 2006

[79] A.M. Kaheel, H. Alnuweiri and F. Gebali, "A New Analytical Model for Computing Blocking Probability in Optical Burst Switching Networks", IEEE Journal on Selected Areas in Communications, Vol. 24, No. 12, pp. 120 - 128, 2006

[80] Z. Lu, Z. Xu, B. Wan, M. Zhang and P. Ye, "Performance Analysis of Burst Assembly under Self-Similar Traffic with Measured WAN Packet Size Distribution", in Proceedings of International Conference on Communications and Networking in China, 2006

[81] X. Lu and B. L. Mark, "Performance Modelling of Optical-Burst Switching with Fiber Delay Lines", IEEE Transactions on Communcations, Vol. 52, No.12, pp. 2175-2183, 2004

[82] M. R. Garey and D. S. Johnson, "Computers and Intractability: A Guide to the Theory of NP-Completeness", W.H. Freeman, 1979.

[83] V.N. Hsu, T.J. Lowe and A. Tamir, "Structured P-facility Location Problems on the Line Solvable in Polynomial Time", Operations Research Letters,Vol. 21, pp. 159–164, 1997.

[84] G. Apostolopoulos et al, "QoS Routing Mechanisms and OSPF Extensions", RFC 2676, August, 1999

[85] S. Shenker, “Fundamental design issues for the future Internet”, IEEE Journal on Selected Areas in Communications, Vol. 13, No. 7, pp. 1176 - 1188, 1995