DISSERTATION

Analysis of Optical Burst Switched Networks with Edge-Core Joint Nodes

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften unter der Leitung von

O. Univ. Prof. Dipl.-Ing. Dr. techn. Harmen R. van As
E389 – Institute of Telecommunications
Technische Universität Wien

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von

Farrukh Zeeshan Khan
Matr.-Nr. 0728661

Wien, im June 2012

بسم الله الرحمن الرحيم

Zusammenfassung

Die Entwicklung von Hochleistungsübertragungstechnologien wird vom kontinuierlichen Wachstum des Datenverkehrs vorangetrieben. Von der dritten Generation der optischen Netze wird erwartet, dass deren photonische Verschaltung die zukünftig benötigte Bandbreite und Skalierbarkeit bereitstellt. Optische Burst Vermittlung (OBS) wird als Vermittlungsprinzip präsentiert, das durch statistisches Multiplexen in der optischen Domäne beides, Flexibilität und Effizienz, bietet.

Ein OBS Netz kann in Randknoten und Kernknoten unterteilt werden. Digitale Datenpakete werden in Randknoten zu Bursts zusammengefasst, und vice versa extrahiert. Somit verbinden Randknoten die digitale mit der optischen Domäne. Kernknoten leiten die optischen Bursts wellenlängengemultiplext über Glasfasern Schritt für Schritt in Richtung ihres Zieles, sind also photonische Vermittlungsknoten. Dabei wird im zumeist ein Reservierungsprotokoll ohne Bestätigung verwendet. Die Übertragung vieler Datenpakete in wenigen Bursts reduziert den Verarbeitungsaufwand je Paket, und die unbestätigte Reservierung ermöglicht Latenzen die wesentlich kleiner sind als jene die für verlässliche Durchschalteverbindungen erforderlich wären.

In der Praxis werden die meisten OBS Knoten sowohl die Funktionalität von Randknoten als auch jene von Kernknoten bereitstellen müssen. Diese Dissertation befasst sich mit der Studie dieser kombinierten Knoten, im weiteren als ECJN bezeichnet. Eine wesentliche Herausforderung beim Entwurf eines ECJN ist die faire Bedienung von lokal erzeugten Bursts und weiterzuleitenden Bursts je Ausgangskanal. Erstere können warten, zweitere nicht, womit sich ein hybrides Warteschlangen/Blockierungs-System ergibt. In dieser Dissertation werden die Eigenschaften von aus ECJN bestehenden OBS Netzen mittels Simulationsstudien evaluiert. Es wird gezeigt, dass der weiterzuleitende Datenverkehr die Wartezeit der lokal einzufügenden Burst beeinflusst und vice versa, dass die Verlustrate der weitergeleiteten Bursts durch den lokal eingefügten Datenverkehr erhöht wird.

Die Kombination von speicherlosen Kernknoten und unbestätigter Reservierung führt dazu, dass gelegentlich ein Burst nicht weitergeleitet werden kann und verworfen werden muss. Folglich ist ein garantierte Vermittlung

für OBS Netze dieser Art nicht möglich. Um diese Verluste gering zu halten wurden etliche Staubeseitigungsansätze in der Literatur vorgeschlagen: Glasfaserverzögerungsleitungen und Umleiten sind wohl die am allgemeinsten anwendbaren.

In Bezug auf ECJN erkennen wir, dass alle Funktionen für das elektronisches Speichern und spätere Weiterleiten von blockierten Bursts vorhanden wären. Im zweiten Teil der Dissertation wird die Speicherung blockierter Bursts als Staubeseitigungsmechanismus untersucht. Da elektronisches Speichern dem Grundgedanken von OBS widerspricht, konzentrieren wir uns auf limitierten Einsatz dieser Option (RIB) und zeigen mittels Simulation, dass bereits mit geringem Zwischenspeichern die Burstverlustrate deutlich reduziert werden kann.

Im letzten Teil beschäftigt sich die Dissertation mit adaptivem Routing. In der Literatur vorgeschlagene Strategien werden mittels Simulationsstudien untersucht und verglichen. Ein smartes Routenselektionsschema das sich unterschiedlichen Netztopologien anpasst wird definiert, und gezeigt, dass dieses in allen untersuchten Szenarios von Vorteil ist. Sowohl für gleichverteilte Last als auch distanzabhängige Verkehrsverteilung wird damit die Burstverlustrate verringert und eine höhere Netznutzung erreicht.

Abstract

The evolution of high speed transmission technology is driven by the continuous growth of data-traffic. Third generation optical networks are envisioned as the most appropriate solution to meet future bandwidth and scalability demands by means of all-optical switching. Optical burst switching (OBS) is proposed as a switching paradigm that offers both, flexibility and efficiency, through exploiting statistical multiplexing in the optical domain.

The OBS networks may be divided into edge nodes connecting the optical domain with the digital domain, and core nodes that switch optical bursts toward their destination. Incoming data packets are assembled into optical bursts at ingress nodes. These bursts are then transmitted hop-by-hop over wavelength division multiplexed (WDM) optical fiber links, commonly using an unacknowledged one-way reservation protocol. The transmission of many packets encapsulated in few bursts reduces the per-packet processing overhead, and the one way reservation scheme enables a latency considerably shorter than the latency required to reliably setup an optical circuit.

For practical deployments the majority of OBS nodes will have to combine both functionalities, that of edge nodes and that of core nodes. This thesis focuses on the study of these combined nodes, hereinafter called edge-core joint nodes (ECJN). The major challenge in the design of an ECJN is that the locally assembled bursts and the transit bursts are switched to the same output channels. The first can be buffered, the latter not, which causes a hybrid queueing/loss system. In this PhD thesis, the performance of OBS networks composed of ECJN is evaluated through simulation studies. It is shown that the transit traffic volume affects the waiting time of locally assembled bursts, and vice versa, that a high ingress data-rate causes increased burst loss-rates for the transiting bursts.

The combination of bufferless core nodes with a one-way reservation strategy causes that sometimes a burst would require a currently occupied resource, and that in consequence it is dropped. Thus, assured transmission is out of the scope for any OBS network of this kind. To reduce this issue several contention resolution approaches have been proposed in the literature; fiber delay lines (FDLs) and deflection routing are the most generic among them.

Considering ECJN we recognize that these in principle have all the facilities required to electronically buffer and later re-insert a blocked transit burst. In the second part of this PhD thesis the option to buffer transit bursts is examined as contention resolution approach. Because electronic buffering is opposing the OBS intention, we concentrate on restricted intermediate buffering (RIB) and show by simulation results that with limited buffering options the burst loss rate can be significantly reduced.

Finally, the PhD thesis emphasizes on adaptive burst routing techniques. Strategies proposed for OBS in the literature are evaluated and compared by means of simulation studies. A smart route selection strategy is defined, one which is adaptable to different network topologies, and it is shown to perform better for all considered scenarios. Also for uniform as well as distance dependent traffic distribution it achieves reduced burst blocking probability and better network utilization.

Acknowledgements

First of all, I thank Allah, the almighty, for giving me enough strength to complete this thesis and for blessing me with valuable people who supported me throughout the completion of this work. I owe special thanks to my supervisor Prof. Dr. Harmen R. van As for accepting me as his student in the first place, for his faith in me during PhD, and for providing me direction and invaluable assistance. He has always encouraged me to consider the generality and big picture of my work.

I am also highly thankful to all of my colleagues in the group for their consistent support from beginning to end. I am grateful to all of them for making my life easier in Austria. I owe special thanks to Muhammad Faisal Hayat and Gerald Franzl for their always available technical support. I am also thankful to Christoph Egger and Lukas Wallentin for their support during the development of simulator.

I would also like to thank Higher Education Commission (HEC) of Pakistan for awarding me PhD scholarship. I would like to thank all my friends in Vienna with whom I spent memorable time.

I am grateful to my wife, Ramla, for her infinite patience during my studies. And last, but most important, mine deepest and warmest gratitude goes to my siblings, Faiza, Asad and Aneeza who have always supported and encouraged me along my way. I especially want to send my deepest love to my mother and father, for being too important to me and it is you to whom I dedicate this thesis.

Vienna, 2012 Farrukh Zeeshan Khan

Contents

Chapter 1

Introduction

The exponential growth of internet traffic is actuating the need for high speed transmission technology that can meet the bandwidth demands of todays bandwidth-intensive applications and society's increasing reliance on communication networks. Wavelength division multiplexing (WDM) networks with capacity to transmit several terabits of data over optical fibers, offered huge capacities and proved to be the most promising transmission technology in the last decade. In the deployment of WDM networks, the optical signal has to go through a bottleneck of optical-electrical-optical (OEO) conversion at the core part of the network. The next generation of optical networks are supposed to provide the facility of optical switching, with the breakthrough in optical technology, where the optical signal travels all optically through intermediate nodes, resulting prompt availability of bandwidth to the end users.

Optical circuit switching (OCS), optical packet switching (OPS) and optical burst switching (OBS) are considered to be the important switching techniques available for the deployment of all-optical WDM networks. Optical circuit switching has the limitations of long setup delay, low flexibility, low network utilization, and therefore, can not handle the bursty internet traffic. For optical packet switching, only preliminary components and subsystems are available and this technology is waiting for a technological breakthrough. Therefore, optical burst switching is supposed to achieve the benefits of OCS while avoiding the limitations of OPS, capable to utilize the huge bandwidth of fiber links more efficiently.

1.1 Optical Switching Techniques

The electrical to optical shift of switching technology is stepping through several phases. The first proposed idea was adopted from traditional voice communication circuit-switched networks [Chlamtac92]. Optical circuit switching (OCS) deploys wavelength routing (WR) which offers circuit switching services at the granularity level of wavelengths. The success of electronic data packet networks is supposed to be a driving force for the research in the area of optical packet switching (OPS) based optical networks [Chiaroni95]. Due to the lack of optical processing and buffering, OPS is very difficult to be deployed in the near future. Optical burst switching (OBS) was proposed in [Turner99, Qiao99] with the objective to overcome the low flexibility of OCS and technological limitations of OPS. The principle of operation of all three optical switching technoligies is described in the following sections.

1.1.1 Optical Circuit Switching (OCS)

In OCS or wavelength routed networks, an all-optical wavelength path, called lightpath, is established between source and destination by dedicating a wavelength on evey link along the selected path. The data transmission on a lightpath does not require optical-electrical conversion, and the lightpath is released after the data transfer. In a lightpath, the same wavelength on all the links is used in case no wavelength conversion is available on intermediate nodes, reffered to as wavelength continuity constraint. In OCS networks, the smallest switching entity is a wavelength.

On any fiber link of a wavelength routed network, no wavelength sharing is allowed between two distinct lightpaths. Moreover, lightpath connections are static and may not be able to accomodate the bursty internet traffic in an efficient manner.

1.1.2 Optical Packet Switching (OPS)

In optical packet switching networks, data packets are statistically multiplexed in all-optical domain, with wavelengths sharing on optical links. Packet headers are supposed to be optically processed, but due to hardware constraints, packets are delayed (temporarily buffered) at the switch input and packet headers are processed in the electronic domain. Packets are processed and forwarded hop-by-hop until they reach at the destination. OPS is foreseen as a promising technique for the next generation WDM networks.

1.1.3 Optical Burst Switching (OBS)

OBS was proposed to overcome the shortcomings of both OCS and OPS. In OBS networks, the data packets are assembled into a super packet, called burst. A control packet is sent to the core network, which reserves the resources for the incoming burst. The burst cuts through the core network all-optically up to the egress node, which disassembles the burst into packets and forwards the packets to the destination networks. OBS provides better network utilization as compared to OCS networks, and technologically less complex than OPS networks.

1.2 Focus of the Thesis

Optical burst switching (OBS) is the most promising optical switching technology in terms of architecture, control coplexity, performance and flexibility. The motivation behind this work is in depth analysis of important functions at node and network level.

OBS networks are divided into two functional domains, the edge and the core. The edge node is the interface of the core optical part, where IP packets arrive from multiple sources. Packets for the same destination egress nodes are assembled into one single unit, called burst. After the completion of the burst assembly, a control packet is sent to the selected path to reserve the resources for the considered burst. The control packet goes through the core network hop-by-hop on dedicated control channels, and processed electrically. The core nodes reserve resources for incoming burst, and after an offset time, the edge node releases the burst. The burst travels all-optically to the destination egress edge node which disassembles the burst into packets and forwards them to corresponding destinations.

A vast literature is devoted to the study of edge and core nodes, and their individual components. In future deployments of dynamically reconfigureable networks, most of the nodes have to combine both the functionalities of edge nodes and core nodes. In this thesis, combination of both edge and core node, called edge-core joint node (ECJN) is studied, individual component functionalities for mixing of local assembled and transit traffic is studied.

In general, OBS networks use one way reservation schemes, resulting an end-to-end transparent connections. Due to the statistical multiplexing, some bursts competing for the same resources contend with each other, and some requests are dropped due to the unavailability of resources. The contention in the core network and retransmission of data from upper layers causes congestion in the core network, which significantly degrades the performance of OBS network. Therefore, contention resolution mechanisms are required to resolve contention.

The architecture of an ECJN provides the facility to electronically buffer the transit bursts, which were going to be dropped due to the unavailability of output wavelength channel. Intermediate buffering using ECJN is proposed as a contention resolution mechanism, which also improves the fairness among the drop rate of bursts with different hop counts.

One of the proactive contention resolution strategy is adaptive routing, which balances the load on the network and better utilizes the network resources. Most of this work is devoted to the study of suitable routing strategies for OBS networks to reduce the possibilities of contention. It is the trend set in OBS studies that newly proposed route calculation and selection strategies are compared with the shortest path routing. The reason behind this is unavailability of complete simulation parameters, and custom development of simulation environments. In this work, some routing strategies are selected from the literature and simulated under uniform network scenarios for fair comparison. At the end, a new route selection strategy is proposed which gives better results in terms of the burst loss rate.

1.3 Contributions of the Thesis

Main contributions of the thesis which address the major issues in the development of OBS networks are:

1. Development of an OBS simulator as an extension to the existing inhouse developed simulator, IBKSim. The OBS part of the IBKSim is capable of simulating individual nodes as well as complete OBS networks. Simulator's modular structure provides high degree of flexibility and its object oriented design approach facilitates further extension of components and modules.

2. Edge-core joint node (ECJN) is studied for the future deployment of OBS networks, its architecture and individual components are analyzed.

3. A new scheduling algorithm, called here the composite edge core scheduling with void filling (CECS-VF) is proposed for the mixing of local and transit traffic in an edge-core joint node of OBS network.

4. Restricted intermediate buffering (RIB) is proposed as a new contention resolution approach where contending bursts may be buffered in the core part of the OBS networks. It is shown that the intermediate buffering also helps to improve the fairness in terms of drop rate of different hop-count bursts.

5. Adaptive routing in OBS networks is studied, and routing strategies from the literature are simulated under uniform network parameters to compare the results. A new route selection strategy is proposed which is adaptable to various network topologies and performs better in all considered scenarios. The results under uniform and distance dependent traffic scenarios are compared with the selected routing strategies to show the improvement in terms of burst blocking probability and better network utilization.

1.4 Organization of the Thesis

This thesis consists of six chapters. **Chapter 2** summarizes the fundamentel concepts of OBS networks such as burst assembly, routing, channel scheduling, signaling and resource reservation schemes, contention management and QoS support in OBS networks.

Chapter 3 gives a brief introdution of the OBS simulation environment developed to assist the study of OBS networks. The edge node, core node and edge-core joint node with respect to the internal node structure and linking of components are discussed. A high level view of implemented algorithms is also presented.

Chapter 4 consists of detail description of an edge-core joint node (ECJN), requirements of such a node, its architecture and functionalities. The scheduling and routing of the joint node is a combination of both edge and core nodes. As for both functionalities, a node has to deal with two types of traffic. It is shown that higher transit traffic effects the local traffic, and vice versa, the high local load increases the drop rate of the transit traffic. Restricted intermediate buffering is proposed for contention resolution in the core part of the network, simulation results prove that the use of intermediate buffering reduces the overall loss rate.

Chapter 5 discusses the adaptive routing in OBS networks. Existing routing strategies from the literature are implemented in simulation and the results are compared for three network topologies. Results show that different routing strategies give different results in different network topologies. A new route selection strategy is proposed which performs better in comparison to all the considered topologies.

Chapter 6 concludes the thesis.

Chapter 2

Optical Burst Switched Networks

Optical burst switched (OBS) networks were proposed to overcome the disadvantages of optical circuit switching (OCS) and hardware limitations of optical packet switching (OPS). The main features of OBS networks are aggregation of data packets at the edge nodes, strict seperation of control and data, electrical control information processing, and all-optical data transmission without optical-electrical conversion. In the last decade, a lot of research is devoted to analyze the functionalities of OBS networks which include burst assembly, resource reservation, routing in OBS networks, wavelength scheduling, offset time issues, quality of service and contention resolution approaches.

This chapter describes the architecture of OBS network. Its two fucntional blocks, the core and the edge are discussed and proposed strategies for individual functions of OBS network are described in detail. Section 2.2 and Section 2.3 provide the architecture of edge and core node. Section 2.4 explains the assembly process and proposed assembly algorithms respectively. Section 2.5 highlights the routing and load balancing in OBS networks. Section 2.6 explains different scheduling algorithms proposed in the literature. In Section 2.7, different signaling and resource reservation schemes in OBS networks are discussed. Section 2.8 is dedicated to the contention resolution approaches proposed in the literature. Section 2.9 gives the quality of service support in OBS networks.

2.1 OBS Network Architecture

Optical burst switched (OBS) networks are divided into two functional domains. Ingress and egress nodes are termed as edge part of the network comprising of edge nodes, whereas, the intermediate or core part of the network consists of core nodes. The architecture of the OBS network is shown in Figure 2.1.

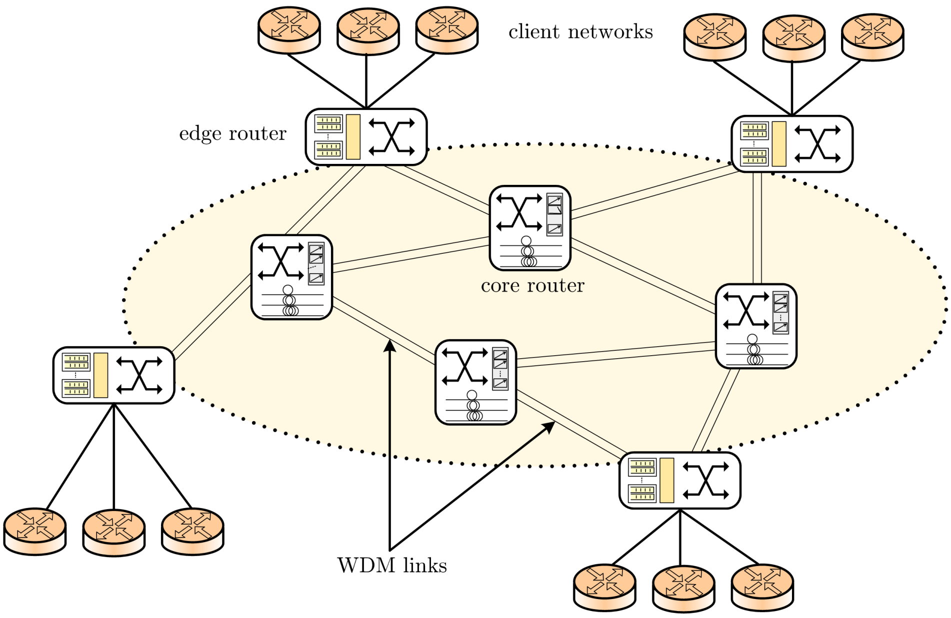


Figure 2.1: Optical burst switched (OBS) network.

The core nodes have WDM connections between each other and to the edge nodes. The edge nodes, connected with client networks such as ethernet, IP networks and MPLS networks, act as an interface of all-optical core network. Data packets from the client networks, destined to the same egress edge node, are aggregated into the bursts and transmitted optically to the core network. A control packet is sent before release of the burst on the selected path of the burst which reserves the resources for the preceding burst. This control packet, referred to as a burst control packet (BCP) is forwarded optically on the dedicated control channels, contains information about the incoming burst such as arrival time, length, wavelength, source, destination, class of service, index of the selected path[1] and other required

[1]In case of adaptive routing, one of the path is selected for every burst from the k

information. With the help of BCP, the core nodes perform burst forwarding and reserve resources for the incoming burst. Burst forwarding is done after routing decision in which output link of the core node is determined for the destination of the burst. Resource reservation function reserves the output wavelength channel on the selected output link for incoming burst. In case, at the time of burst arrival, the output wavelength is already occupied, a contention resolution mechanism is applied if applicable. If contention resolution mechanism remains unable to find alternative resources for the burst, the burst is dropped at the current node and BCP is not forwarded to the next nodes. When the burst reaches at the destination edge node, the edge node disassembles the burst and forwards the data packets to destination client networks.

2.2 Edge Node

Edge nodes are connected to client networks which can be any legacy network, for example, IP, MPLS, ATM, SONET/SDH. On the other side, the edge nodes have WDM connections with core nodes. An ingress edge node converts the data packets from client networks in to the OBS packet, called burst. An egress edge node is responsible for retrieving the data back from the burst and forwards it to the respective client networks. Following are the primary functions of an OBS edge node:

- aggregation of the data from client networks
- burst assembly
- route calculation or selection for the burst
- channel scheduling
- offset time management
- generation of a burst control packet
- and burst transmission

As shown in Figure 2.2, packets from client networks arrive at the edge node which are classified according to their forward equivalence class (FEC). Each FEC describes packets of similar characteristics, for example, packet destination, quality of service (QoS), class of service (CoS). Packets for each FEC are aggregated into different queues and bursts are assembled according to the burst assembly algorithm used. Feasible path for the destination of the burst is found and the offset time is computed according to the number of hops the burst has to traverse. Additional offset time

available paths, in this case, the BCP contains the index of the selected path needed to get the next node from the forwarding table.

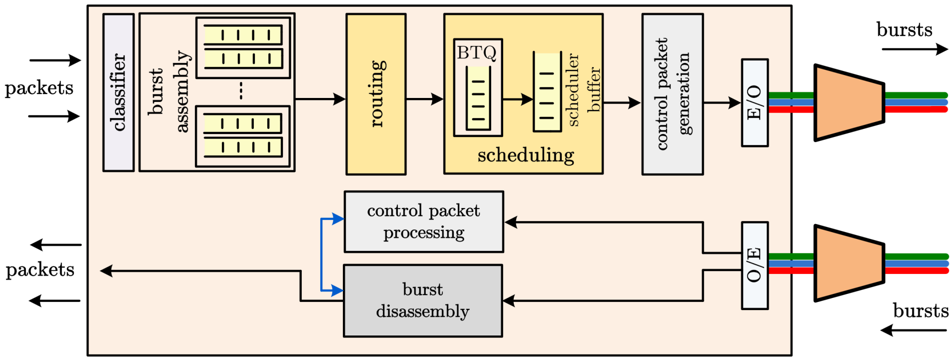


Figure 2.2: OBS edge node architecture.

can be applied in case QoS provisioning is available. The burst is buffered in the burst transmission queue, and scheduling algorithm searches for the output wavelength. If any of the wavelength is available for the time of burst length (according to the scheduling algorithm used), it is reserved for the burst transmission. A burst control packet is sent to reserve the resources for the burst, and the burst is forwarded to the scheduler buffer. The burst is released after the offset time is elapsed. Offset time is the time difference between the release of burst control packet and the corresponding burst. The control packet is processed electronically, for which every core node needs processing time and the setup time for the incoming burst. Therefore, both the control packet and the burst are released with a time difference, which is usually the multiple of processing time of one core node and the number of hops the burst needs to traverse to reach its destination.

2.3 Core Node

Core nodes switch the incoming bursts all-optically from one fiber link to another. The architecture of a core node is given in Figure 2.3. Core node performs following primary functions:

- electronic processing of control packet
- burst forwarding
- scheduling
- contention resolution (if applicable)
- optical switching of burst

In each core node, there is an interface of seperation of control and data wavelengths. The core nodes have two parts, control unit and switch-

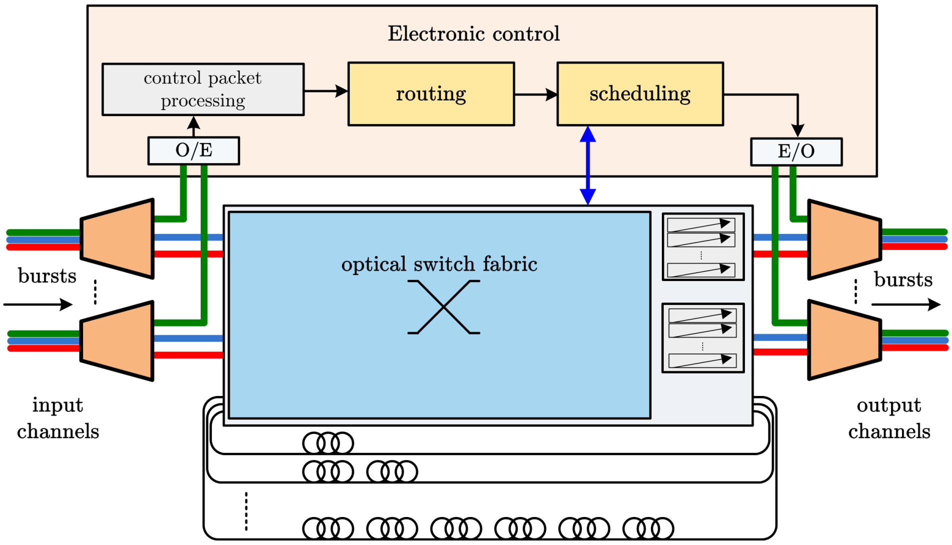


Figure 2.3: OBS core node architecture.

ing unit. Control packets arriving on dedicated control channels, are forwarded to the control unit, converted to electrical domain, and processed to retrieve the required information. Burst forwarding is performed from routing lookup table, and the scheduling module searches for the availability of output wavelength for the incoming burst. If wavelength is found, it is reserved and switching operation is scheduled for the arrival time of the burst, and control packet is transmitted optically to the next node. In case of unavailability of the output wavelength channel, the core node applies the contention resolution mechanisms available in the router, such as, wavelength conversion, delaying using fiber delay lines (FDL), deflection routing, etc., and updates the control packet for the change of resources, and transmits it to the next node. In case of unavailability of resources, even with the contention resolution mechanisms, the burst is dropped at the current node, and the control packet is not forwarded to the next node. The control unit configures the switching unit for the optical switching. Switching unit consists of optical switching cross-connect (OXC) and other dedicated components. The other components in the switching unit are wavelength converters and fiber delay lines which are used for the contention resolution.

2.4 Burst Assembly

The key feature of an edge node is to assemble the packets into a super-packet, known as burst [Qiao99]. Burst assembler performs the burst assembly, in such a way that there exists one queue for each destination, incoming packets are classified according to their destination. If OBS core network supports service differentiation based on the class of service (CoS), than instead of one queue per destination, there may be K queues for each destination, given that there are K supported service classes [Xiong00]. The architecture of a typical burst assembler is shown in Figure 2.4.

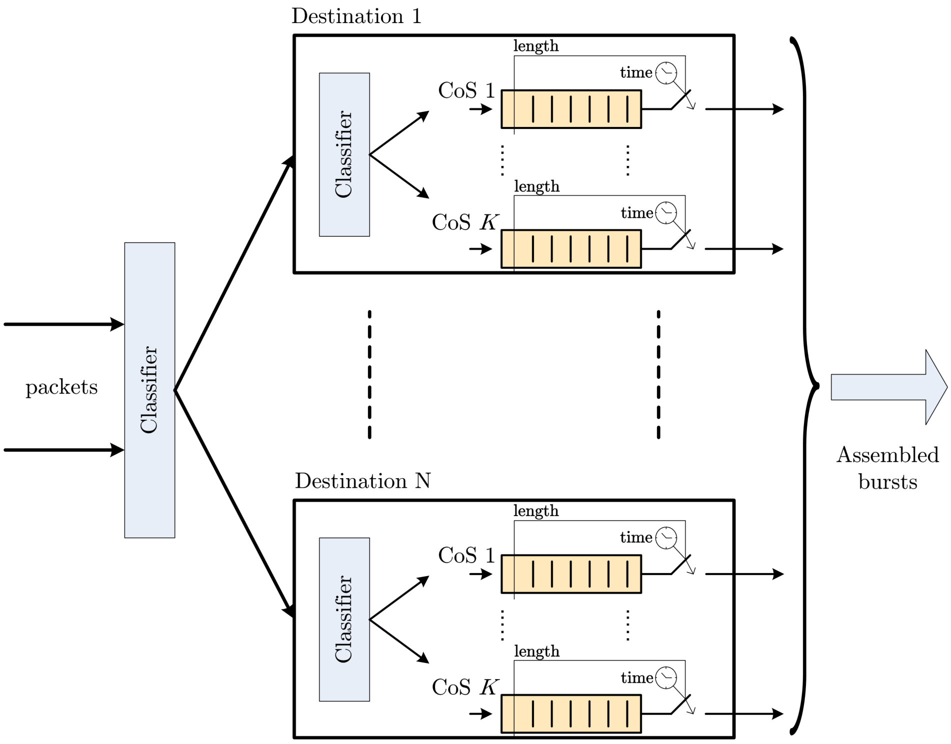


Figure 2.4: OBS assembler architecture.

Packets are aggregated into the specific queues, and when a certain condition is triggered, all the packets in the queue are grouped to form a burst. Burst assembly algorithms define when to complete the burst assembly process. These algorithms can be time based, size based or hybrid based assembly algorithms. In time based assembly algorithms, arrival of first packet in the queue starts the assembly, and the assembly is finished when the time threshold is met. Time based assembly algorithms assemble bursts of large sizes in high load scenarios, and vice versa, small burst sizes in low load scenarios. Whereas, the interdeparture time is deterministic. In the size based assembly algorithms, the burst is assembled when the size of the queue reaches a certain threshold. The size based assembly produces deterministic size bursts while the bursts interdeparture time is low in high loads and high in low loads. Moreover, in low load scenarios, the assembly time increases.

To overcome the shortcomings of time and size based assembly algorithms, hybrid assembly is proposed [Yu02], in which burst assembly finishes when the timer expires or the burst length exceeds the fixed threshold, whichever happens first.

The interdeparture time and the size of the bursts are dependent on the type of assembly scheme adopted. Traffic characteristics of the assembled bursts have been studied in [Yuan11, Hayat11a].

2.5 Routing in OBS Networks

Assembled bursts are forwarded to the routing module, which either calculates the feasible path for the burst, or selects the most feasible path from the pre-calculated paths. Path calculation and selection in OBS netwoks are discussed in detail in Chapter 5. Path calculation or selection is performed only in edge nodes, whereas, only forwarding is performed in the core nodes.

In case adaptive routing is implemented in the network, the core nodes update the status of their output links periodically, based on a certain criteria. These status messages can be probe based or broadcast based. The edge nodes utilize this status information while selecting the path for bursts. The status information can be link load, link utilization, link blocking probability or any other kind of information required by the implemented route selection scheme. Adaptive routing performs better than the shortest path routing, reduces overall burst loss rate, balances load on the links, increases throughput of the network and as a result, better utilizes the network resources.

2.6 Channel Scheduling

After the routing decision in an ingress edge node is made for a burst, the burst is ready for channel scheduling. Similarly, in the core node, after the forwarding decision is made, channel scheduling decides which output wavelength is to be reserved for the considered burst. Scheduling algorithm also undertakes the assignment of wavelength converters and fiber delay lines (FDL) if contention resolution is needed in the core node.

A channel scheduling algorithm keeps reservation record of all the output wavelengths. When a burst or a control packet is arrived, the channel scheduling algorithm finds the time slot of burst size to schedule the burst on a particular wavelength. If there exists no free slot to fit the burst, alternative resources are utilized using contention resolution mechanisms, if applicable in the node.

Channel scheduling algorithms can be categorized into two types, non-void filling and void filling scheduling algorithms. Non-void filling algorithms only keep track of the latest available unscheduled time, which usually is the end time of the last scheduled burst on the wavelength. These algorithms are very simple to implement, but result in low utilization of the channels. Void filling algorithms are taking advantange of the voids created by the random scheduling of bursts and statistical multiplexing using advance reservation. These algorithms keep record of each void, and try to schedule the bursts even between the two already scheduled bursts. Void filling algorithms are reletively complex in implementation but result in high utilization of output channels and perform better in terms of burst blocking probability, end-to-end delay and increase network throughput.

Figure 2.5 describes the scheduling mechanisms for different proposed algorithms. First fit unscheduled channel (FFUC) is a non-void filling scheduling algorithm, which assigns the first available channel to the incoming burst. Latest available unscheduled channel (LAUC) proposed in [Xiong00], is also a non-void filling algorithms. LAUC assigns the latest available channel to the incoming burst, as shown in the Figure 2.5, instead of selecting wl_0, it will select wl_1 for scheduling. The logic behind this is the introduction of minimum gaps, and as a result, it increases the utilization of the link.

First fit with void filling (FF-VF) is a variant of the FFUC algorithm, and uses the voids to the incoming bursts. It assigns the first available void to schedule the burst. Latest available unused channel with void filling (LAUC-VF), proposed in [Xiong00], schedules the bursts into the voids. Similar to the LAUC, this algorithm schedules the burst on the channel creating minimum void with the previous burst on the channel. This algorithm performs better than FF-VF, because of the reason that it will not schedule the burst in the bigger voids, which will remain available for the

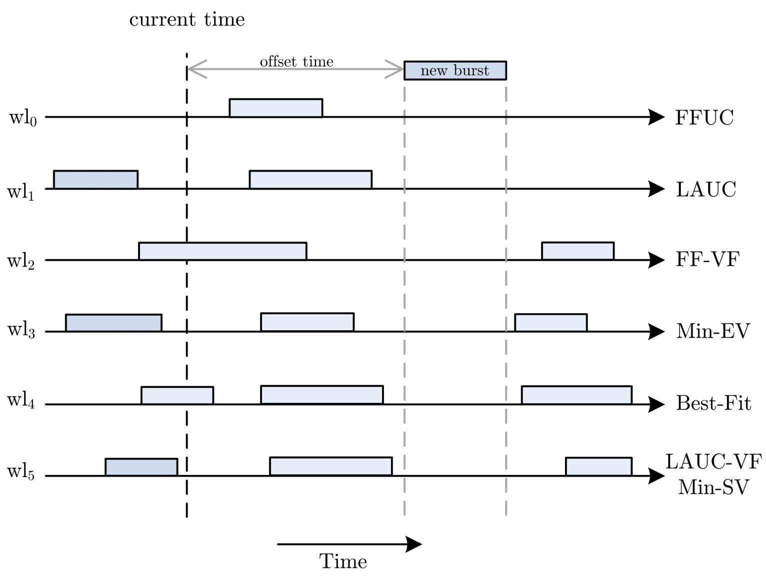
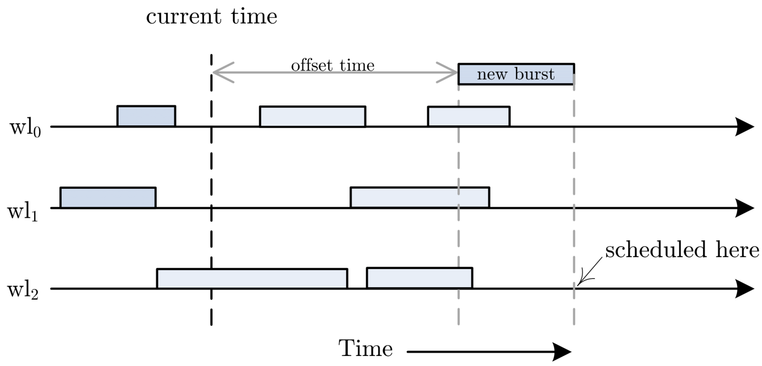


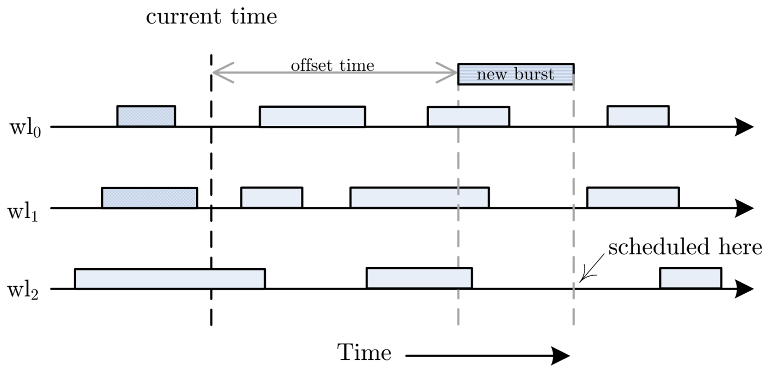
Figure 2.5: Channel scheduling algorithms in OBS networks.

longer bursts. Several variants of LAUC-VF algorithm were proposed in [Xu03], these are, minimum-starting void (Min-SV), minimum-ending void (Min-EV) and best-fit algorithms. All three of them are shown in the Figure 2.5, Min-EV tries to minimize the gap with the next scheduled burst, Min-SV minimizes the gap with the previous burst and Best-Fit searches for the minimum void in total.

If channel scheduling algorithm remains unable to find the available output channel for the incoming burst even after the contention resolution mechanisms, one option is to drop the burst, and most of the scheduling algorithms drop such bursts. Second option is to only drop the portion of the burst which is contending with the already scheduled bursts. This type of dropping is called segmentation based dropping. A scheduling algorithm, minimum overlap channel (MOC) was proposed in [Vokkarane03], which selects the channel on which minimum number of burst segments need to be dropped. An extension MOC-VF was also proposed in the same paper, which schedules the burst in the voids, and selects the channel with minimum possible overlap. Scheduling of MOC is described in Figure 2.6 with and without void filling. This segmentation based dropping improves the link utilization and overall packet drop rate of the network.



(a)



(b)

Figure 2.6: Minimum overlap channel (MOC), (a) without void filling, (b) with void filling.

2.7 Signaling and Resource Reservation

OBS networks can be differenciated from other switching because of its signaling and resource reservation protocols. In an ingress edge node, before transmitting the burst, a burst control packet is transmitted to signal the core nodes for the reservation of resources and the burst is released with the difference of an offset time. Signaling is performed on end-to-end basis for every assembled burst. Two signaling schemes, tell-and-wait (TAW) and tell-and-go (TAG) have been proposed for high speed networks [Widjaja95]. These schemes were originally proposed for ATM block transfer (ABT) protocol and were later adapted for OBS networks [Hudek95].

2.7.1 Tell-and-wait (TAW)

TAW is a two way signaling protocol which performs resource reservation with acknowledgement. As shown in Figure 2.7, the ingress edge node transmits a burst control packet (BCP) to the core nodes, the BCP is processed at every core node, and an acknowledgement finally from the

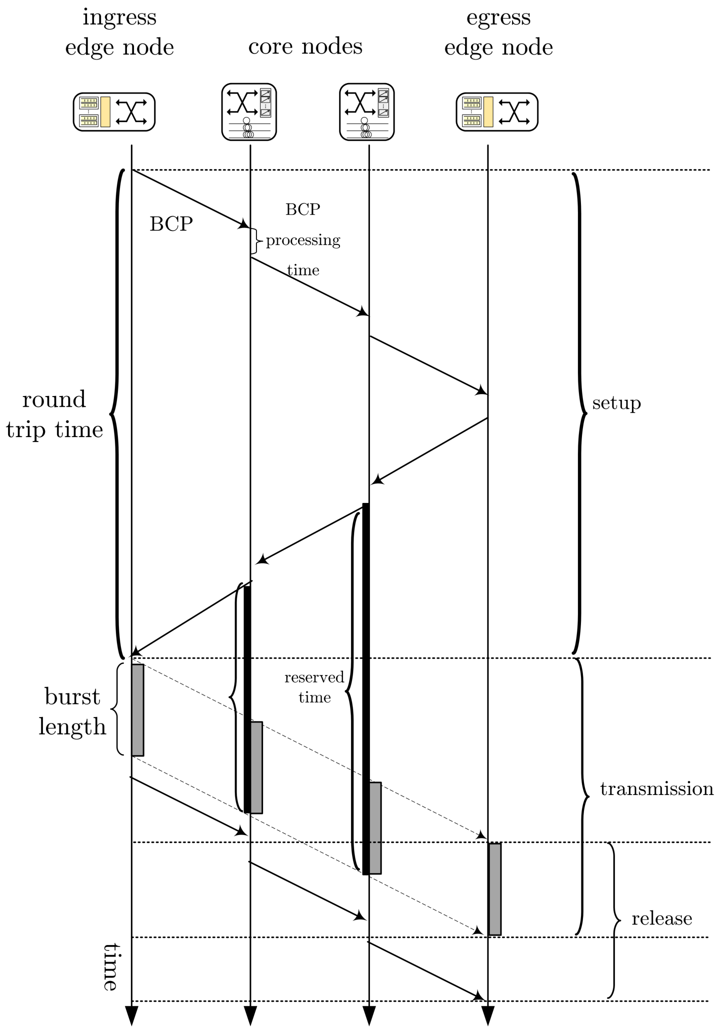


Figure 2.7: Tell and wait (TAW) signaling scheme.

egress edge node comes on the same path. With the acknowledgement, every core node explicitly reserves the resources. The burst is released upon the acknowledgement and a burst release packet (BRP) is necessary for the release of reserved resources. Two-way signaling increases the setup time for the burst transmission, but good for the loss sensitive traffic.

2.7.2 Tell-and-go (TAG)

TAG is one way signaling scheme and does not wait for the acknowledgement after the control packet transmission. The burst is released after a fixed offset time, without knowing that the core nodes remained successful in reservation of resources or not. They caused much reduced setup time, and are very good for delay sensitive traffic. There are two variants of TAG, just-in-time (JIT) and just-enough-time (JET) signaling.

Just-in-time (JIT)

Figure 2.8 describes the functionality of JIT protocol. When BCP reaches a core node, the core node looks for the resources and reserves immediately if

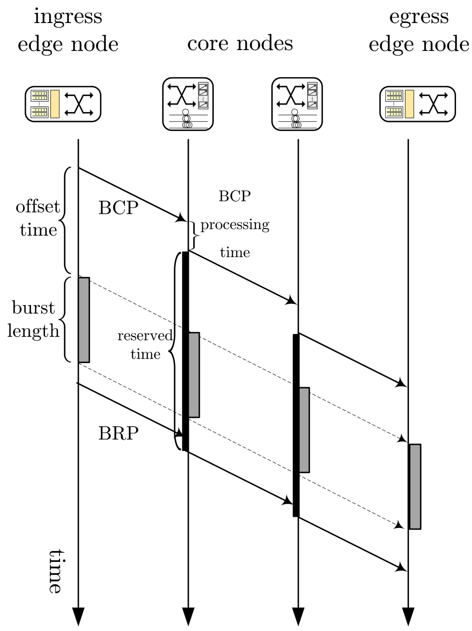


Figure 2.8: Just in time (JIT) signaling scheme. BRP: burst release packet.

available. The burst is transmitted from the ingress edge node after the offset time and after the transmission of complete burst, a burst release packet (BRP) is transmitted for the release of the reserved resources. Therefore, JIT is an immediate reservation and explicit release signaling scheme.

Just-enough-time (JET)

JET was proposed in [Yoo97], performs estimated resource reservation and release. As shown in Figure 2.9, JET reserves the resources from the arrival time of the burst, and they are automatically released after the burst is transmitted. In other words, JET only reserves the resources for the length of the burst in time. In JET the resources are reserved for less time and link utilization is increased with the decrease in signaling overhead.

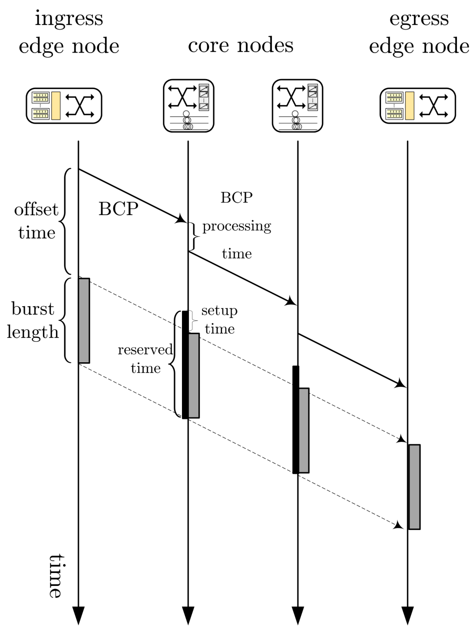


Figure 2.9: Just enough time (JET) signaling scheme.

2.8 Contention Management

In general one way signaling is preferred in OBS networks. When two or more bursts require the same resources at the same time, contention occurs, and due to the bufferless architecture of OBS networks some of the bursts drop in the core network. Dropping a burst containing hundreds of packets, and retransmission of data from upper layers overloads the network. Therefore, the contention resolution is one of the most challenging tasks of OBS networks. A number of contention avoidance and contention resolution mechanisms have been proposed in the literature, which are classified according to Figure 2.10.

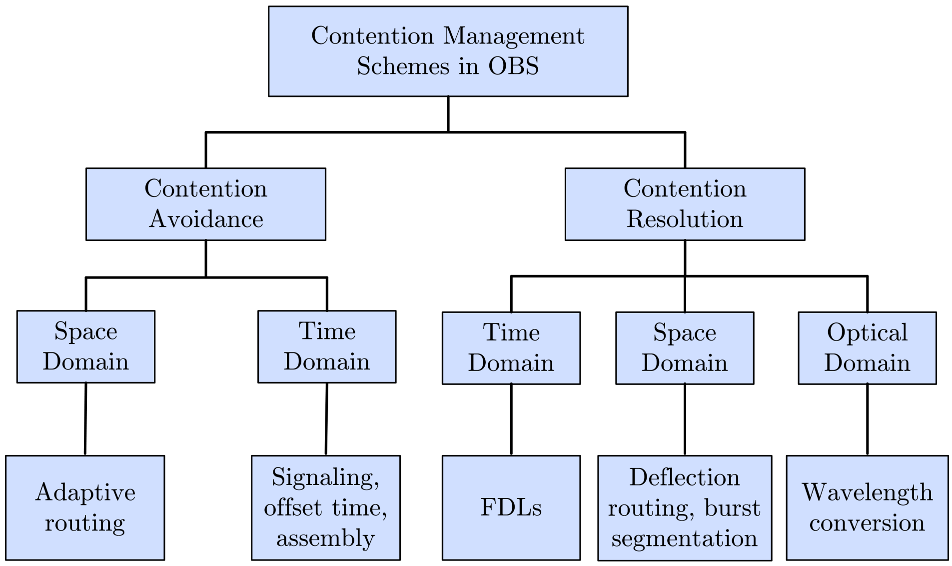


Figure 2.10: Contention management in OBS networks.

There are two main approaches in OBS networks to deal with the contention, contention avoidance and contention resolution:

2.8.1 Contention Avoidance

Contention avoidance is a proactive approach which tries to minimize the contention. In other words, it avoids the contention before its occurance. Adaptive routing is one of the proactive approaches to resolve the contention in space domain. Adaptive routing algorithms are adaptive to the current network situation, for example, link loads, burst losses, link utilization, and select the route with the minimum possibility of contention. That is why adaptive routing algorithms perform better than the shortest path routing in terms of the burst loss probability. Adaptive routing for OBS networks is discussed in detail in Chapter 5.

Two way signaling, offset time selection mechanims and different burst assembly techniques are contention avoidance techniques in time domain.

2.8.2 Contention Resolution

Contention resolution schemes are reactive approaches to resolve the contention. Unlike contention avoidance, these approaches resolve the contention when it occurs. Contention resolution can be categorized into time domain, space domain and optical domain.

Time Domain

Fiber delay lines (FDLs) are proposed as a contention resolution mechanism to delay the contending bursts for a fixed duration of time [Pedro07a]. FDLs are special type of fibers which allow limited fixed delayed transmission of the optical signal. Without any storage, FDLs delay the signal according to the propagation time of the signal in it. The delay achieved is proportional to the length of FDL. For example, a 200-km standard FDL is needed to delay a single burst for 1 ms [Jue05].

FDLs are used in groups to achieve variable delays, called FDL buffers. When the wavelength demanded by the incoming burst is already occupied, the scheduling algorithm looks for the available FDLs in the buffer, and searches for the availability of wavelength after a delay. The wavelength is reserved if it is available after the delay, BCP is updated with the delay and forwarded to the next node.

Space Domain

Contention in space domain is resolved using deflection routing [Wang00], in which the contenting bursts are deflected to less loaded output links. For the application of deflection routing every core node calculates the alternate path for every possible destination. Deflection routing also improves the performance of the OBS network as compared to no deflection routing [Kim02, Lee05]. Deflection routing has the problems of the potential network loops, out-of-sequence delivery of packets and increased end-to-end transmission delay, special algorithms are required to deal these issues.

Burst segmentation [Vokkarane02a, Sarwar08] is another method of contention resolution in space domain. In burst segmentation, only the overlaping segments of the bursts are dropped instead of dropping the whole burst. The burst, at the moment of assembly is divided into segments which are the partitioning points of the burst. Additional control information is attached with each segment. Despite its contribution to the reduction of burst loss rate, burst segmentation adds overhead to the complexity of an OBS network.

Optical Domain

In optical domain, wavelength converters are used to convert the wavelength of a contending burst to another available wavelength. Wavelength conversion is an expensive optical technology, and there exists a variety of types for wavelenth converters. They are mainly categorized as fixed wavelength converters (FWCs) and tunable wavelength converters (TWCs). FWCs have fixed input or output wavelenghts or range of wavelenghts, whereas, TWCs are tunable to a range of wavelenghts both for input and output. The third category, and the most expensive one is capable of converting any input wavelength to any output wavelenght. A lot of research is dedicated to the sharing of wavelenght converters among output channels and output links [Eramo03, Hayat11b].

2.9 QoS Support in OBS Networks

Due to the bufferless nature of OBS networks, the loss in the core network is unavoidable. Therefore, it is very difficult to ensure an absolute QoS in OBS networks. In addition to the burst loss, end-to-end delay is another important performance metric in OBS. End-to-end delay is dominated by the assembly delay, but almost bufferless and fast transmission may result in lower latency than in traditional data networks [Klinkowski07b]. Therefore, relative QoS provisioning is proposed in OBS networks taking into account the burst loss rate of different classes of traffic.

An offset time-based QoS provisioning mechanism was proposed in [Yoo00]. This mechanism assigns an extra offset to the high priority bursts, with a small multiple of low priority bursts' offset, as a result, high priority bursts get less possibility of contending with other bursts. The authors claim that an extra offset time equal to three and five times the mean burst durations guarantees ninety-five percent to ninety-nine percent degree of class isolation respectively.

Different burst assembly mechanisms are proposed to provide QoS in OBS networks. A length-based burst assembly scheme in conjunction with a burst segmentation policy to provide QoS is proposed in [Vokkarane02b].

Different size of bursts are assembled on the basis of QoS differentiation and low burst loss of high priority traffic is achieved. Another two class system with the high and low priority traffic was proposed in [Klinkowski05], in which service differentiation can be provided on the basis of burst lengths when combined with offset time and preemptive priority based QoS mechanisms. Use of void filling scheduling algorithms favors the bursts with shorter lengths, therefore, high priority bursts are assembled with small lengths.

Burst preemption based QoS mechanisms provide QoS to the high priority bursts by preempting the already scheduled low priority bursts. Preemptive latest available unused channel with void filling (PLAUC-VF), proposed in [Kaheel03], guarantees QoS to high priority bursts. Probabilistic preemptive burst segmentation (PPBS), proposed in [Tan04], enables the high priority bursts to preempt and segment low priority bursts in a probabilistic fashion.

Chapter 3

Simulation Environment

This chapter presents an overview of OBS simulator designed to facilitate the study of OBS networks. The general network simulator IKNSim was originally developed in 2004 [Schuringa04], many protocols were added over the years to aid in the research work. The current version of the simulator is refered to as IBKSim. This simulator is capable of simulating tiny queueing models, single network nodes and able to asses the performance of bigger networks. Its modular structure provides high degree of flexibility and object oriented design approach facilitates extension on the development part. A detailed description about the simulator is beyond the scope of this document, this chapter covers only fundamental concepts needed to simulate OBS networks. For the basic concept of IBKSim and the use of XML as configuration and logging language the readers are refered to [Wallentin10].

Later in this chapter the simulation of OBS networks using IBKSim is discussed for edge, core and joint node. Burst assembly, scheduling, routing and statistics in the simulator are discribed. Simulator is validated with comparison of results with Erlang-B loss formula, and at the end some of the simulations results are presented.

3.1 Simulation of OBS networks

The architecture of optical burst switched (OBS) networks has complex structure and components of OBS network have high degree of freedom in traffic pattern, arriving packet distribution, assembly techniques, class of service, routing protocols, scheduling algorithms, buffering requirements, burst size, offset time settings and signaling protocols. So far, a vast literature on OBS is there and analytical modeling providing insights to individual components of OBS networks, but might scarcely cope with multiple factors that may arise in complete network scenarios [Pedrola10]. Therefore, there coexists a need of simulation tools to evaluate the performance of complex network architectures such as OBS.

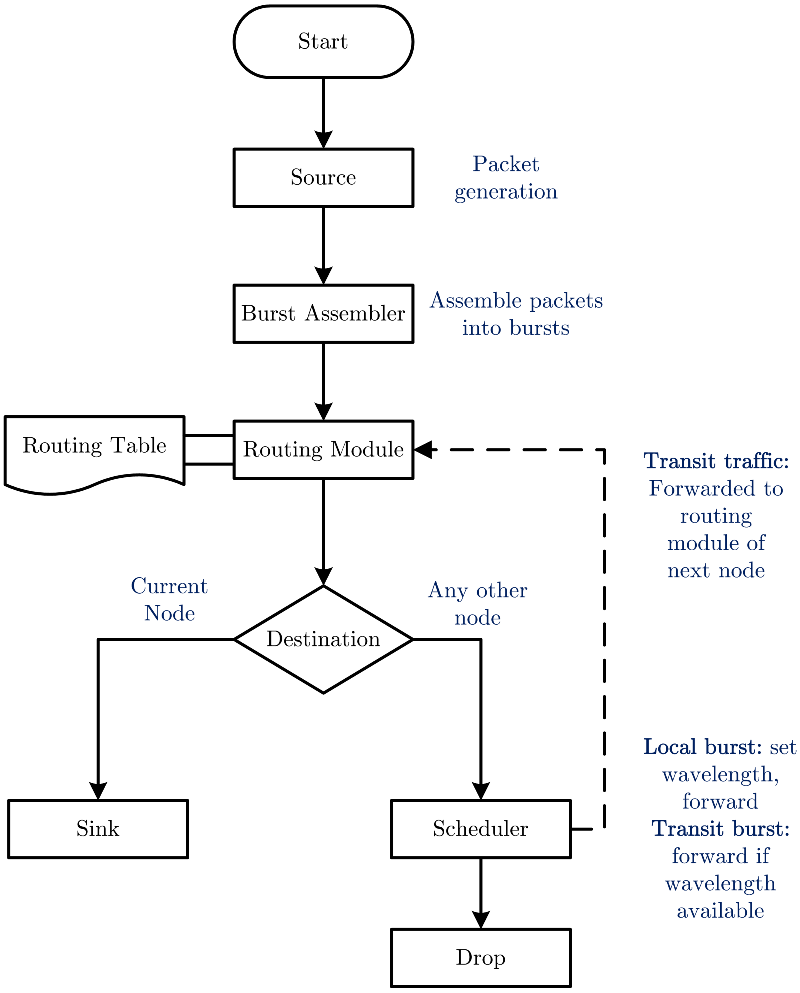


Figure 3.1: Flow chart of a typical OBS simulation.

The OBS module of IBKSim takes XML configuration file as input and generates logging file in the form of XML file. In the configuration file the user can describe the whole simulation scenario, such as, run time of simulation or number of specific events the simulation has to run, network, nodes, components within the nodes, links between nodes and internal links of components within a node. Figure 3.1 demonstrates a typical simulation flow of OBS network.

After the start of simulation, packet sources connected to edge nodes start generating packets using packet size distribution and generate rate given in input configuration file. Packets are forwarded to burst assembler module which assembles packets into bursts according to given assembly scheme. Assembled bursts are forwarded to routing module, which looks up the routing table to forward the burst to relevant scheduler or the sink in case current node is destination for the burst. Scheduler module searches for the output wavelength, if found, the burst is forwarded to the routing module of next node. Simulation lasts until the input simulation time or the number of time the specific event is occured.

OBS is mainly composed of edge router and core router, which further contain many components having number of internal connections. Both edge and core router, and main components are discussed below.

3.1.1 Edge Node

The edge router or edge node of OBS network is responsible for burst assembly, path selection or calculation for the assembled bursts, offset time calculation according to class of service (CoS) and number of hops the burst has to traverse and scheduling of bursts on the output wavelength channels. Figure 3.2 shows the components and internal links of the components within an edge node.

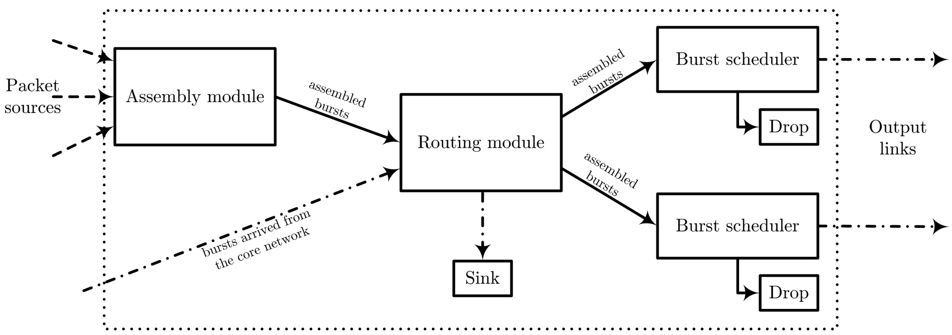


Figure 3.2: Visualization of an edge node and components' links.

Packet sources are connected to burst assembly module discussed in detail in Section 3.1.4, which assembles packets into bursts and forwards them to the routing module. Routing module finds the path for the given destination and forwards the burst to the output link connected to the next node in the selected path for the burst. Each link has a scheduler, and routing module is connected to the scheduler, which searches for the time slot to forward the arrived burst and sets the wavelength of the bursts. The schedulers are connected to the routing module of the next core router or edge router depending upon the topology used. In case the buffer in the scheduler module gets full, the locally assembled bursts are dropped. Similarly, output links of the core nodes are connected to the routing module of next core node or egress edge router. The bursts arriving from the core network are forwarded to the routing module, which are sinked in the current node and statistics are updated.

3.1.2 Core Node

The core router in an OBS network forwards the transit traffic to the next core router or egress edge router all optically. In simulation, core nodes perform similar to the edge nodes except the burst assembly. Moreover, core nodes are also not the destination node of any bursts, therefore, there is no sink in the core part. Components of the core node and their internal connections are shown in Figure 3.3.

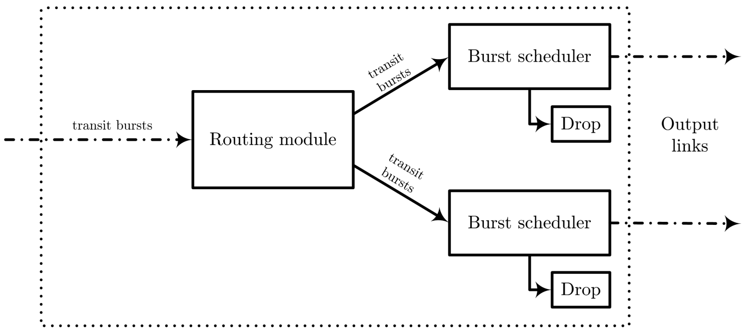


Figure 3.3: Visualization of a core node and components' links.

Transit bursts arrive to routing module of the core router, which selects the next node for this burst from the routing table, subtracts the control packet processing time from the burst's offset time and forwards it to the scheduler module. Scheduler module schedules the burst according to the selected scheduling algorithm, and drops the burst if channel is already occupied or the burst is dropped. In case the scheduling is successful, the bursts is forwarded to the routing module of next node after the propagation delay, computed according to the output link speed and the distance to the next node. The link distance and speed parameters are input by user in the input configuration file.

3.1.3 Joint Edge-Core Node

The joint edge-core node (JECN), described briefly in Chapter 4, assembles the data bursts performing the function of edge router, and forwards transit bursts to the next nodes while performing the function of core router. Components in the simulation node of joint node are shown in Figure 3.4.

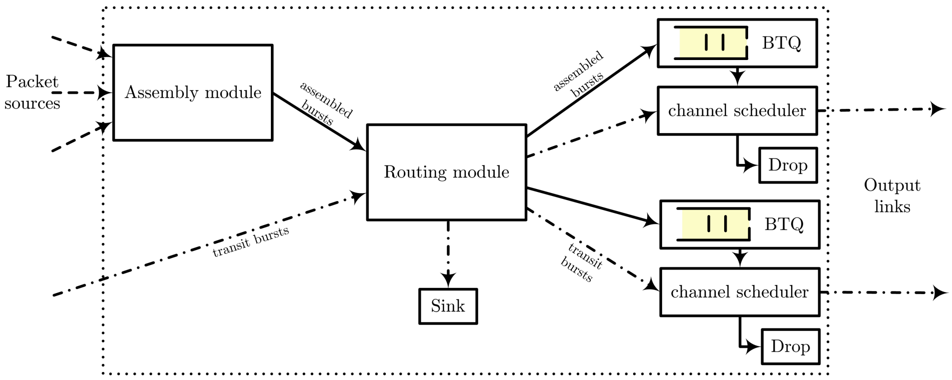


Figure 3.4: Visualization of a joint node and components' links.

Packets arrived from different packet sources are assembled into bursts which are forwarded to the edge scheduler after routing decision by routing module. Transit bursts are recieved by routing module as well, which are also forwarded to scheduler module. The scheduler module of joint node differentiates between locally assembled bursts and transit burst, buffers the local bursts in BTQs, whereas, transit bursts are scheduled without buffering. Forwarded bursts are sent to the routing module of the next node, which after routing decision, forwarded to the channel scheduler.

3.1.4 Burst Assembly

Burst assembly is mainly categorized into three types of algorithms, time based burst assembly [Ge00], size based burst assembly [Vokkarane02b], [Vokkarane02c], and hybrid burst assembly [Yu02]. All three of these algorithms are implemented in simulation, and can be selected via input configuration file. A high level view of burst assembly implementation is given in Figure 3.5.

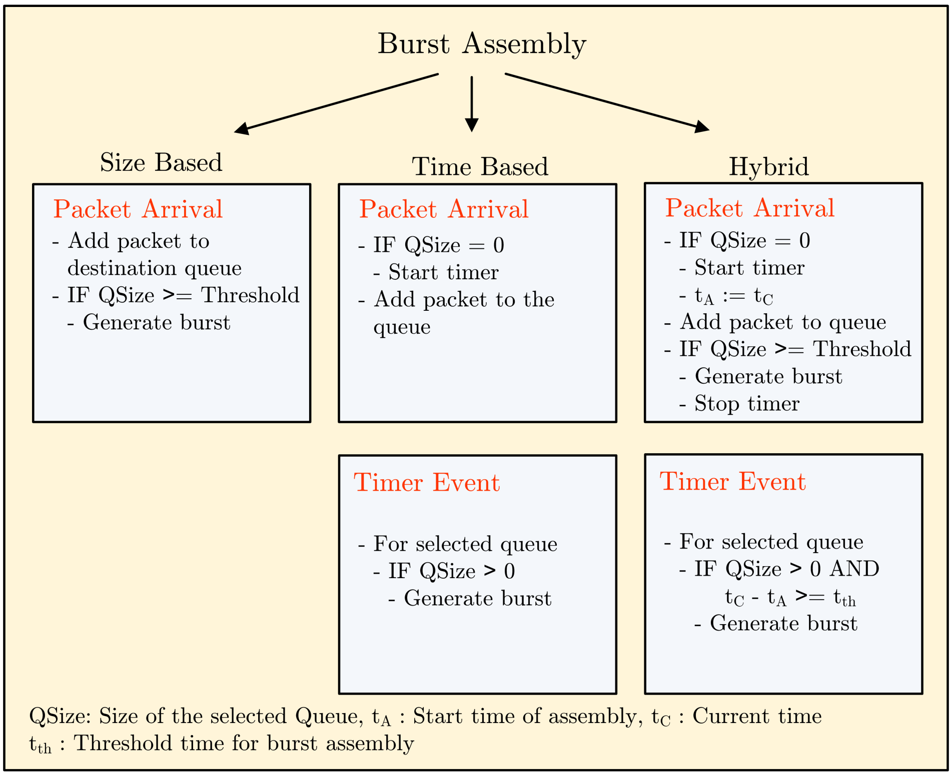


Figure 3.5: High level view of burst assembly algorithm.

In size based burst assembly, there is only one event, i.e., packet arrival. On every packet arrival the packet is added to the specific queue for packet destination. On every arrival, queue size is checked, if it has met the threshold given, the burst is assembled and forwarded to the routing module. In case of time based assembly, on packet arrival event, timer is started if arrived packet is the first packet in the queue. On occurance of timer event, the burst is generated. In case of hybrid burst assembly both size and time based assembly algorithms run in parallel. If packet arrival is the first packet in the queue, the timer is started, and if queue reaches to the size threshold due to addition of current packet, the burst is generated and timer is stopped. When timer event occurs, the algorithm verifies that current time time threshold t_{th} is met, the burst is generated.

Time and size threshold are set in the input configuration file and statistics of burst assembly are updated on completion of burst assembly. Mean number of packets in the burst, average burst size, assembly delay and inter departure time of the bursts are the statistics which can be logged into the output file. The user has freedom of choice in selection of time and size threshold and the selection of assembly algorithm.

3.1.5 Burst Scheduling

Channel scheduling algorithms can be classified into two categories; without void filling and with void filling. For the case of without void filling, latest available unscheduled channel (LAUC) proposed in [Xiong99] is implemented, and for void filling, latest available unused channel with void filling (LAUC-VF) proposed in [Xiong00] is implemented. Both algorithms work for the edge scheduling where the electronic buffer is available for the bursts, and for the core scheduling where the bursts are dropped in case of unavailability of channel. Moreover, composite edge-core scheduling with void filling (CECS-VF) is proposed and implemented for the scheduling of edge-core joint node, discussed in detail in Section 4.2. A high level veiw of LAUC-VF channel scheduling algorithm for core scheduling is presented in Figure 3.6.

The channel scheduling at core acts as pure loss system, the bursts are dropped if there is no output available. A list is maintained at every channel which is sorted by the start time of the bursts in the list, i.e., t_s. In first step upon arrival of the burst[1] the scheduler removes all the bursts from the list whose end time t_E is less than the current time t_C, because of the reason that these bursts are not part of future reservation and therefore, not needed for void filling. In the second step, the algorithm traverses burst list on each channel meeting the wavelenth requirements, if the list is empty, the channel is selected using LAUC algorithm. In case the list has already some bursts scheduled in advance, the scheduler has to find the voids on the list, for this purpose, the burst is temporariliy inserted into list, and verified that if its not contending with any other burst. This operation is performed on all the channels meeting wavelength requirements, and on all the channels in case the full wavelength conversion is set by the user. If there exist more than one option to schedule the burst, the channel creating

[1] Burst control packet (BCP) arrives at the core scheduler, and the scheduler reserves output wavelength for arriving burst. The simulation is dealing with the control packet, but for simplicity, the term 'burst' is used.

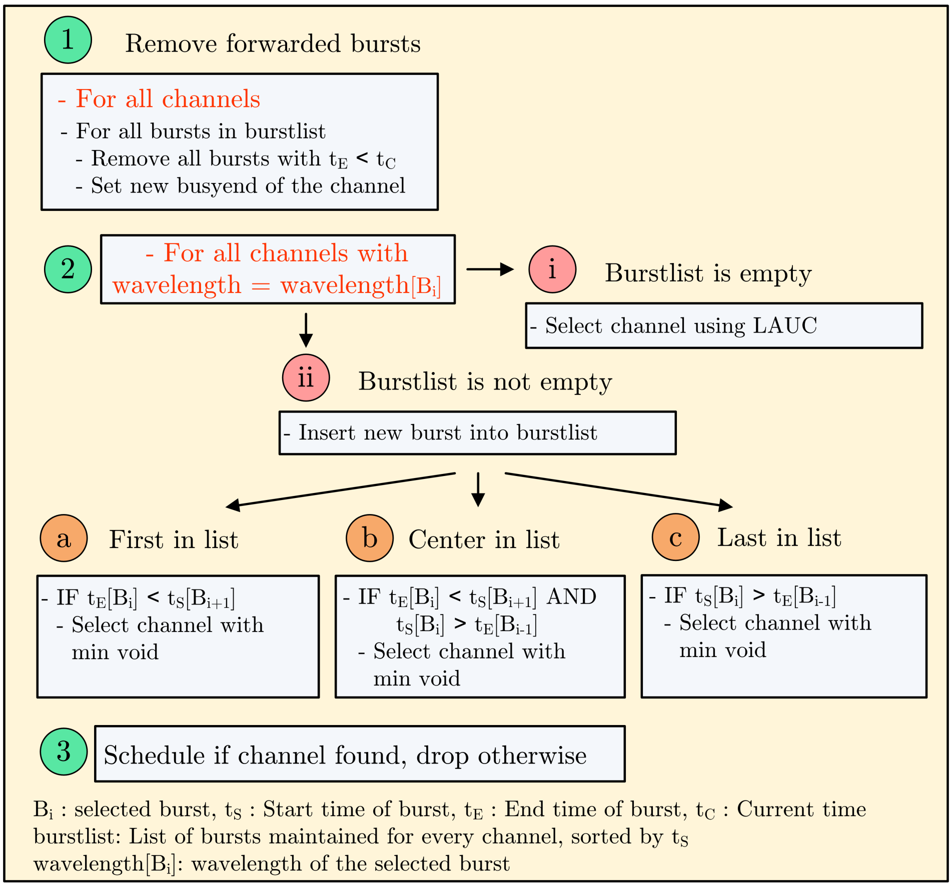


Figure 3.6: High level view of LAUC-VF channel scheduling at core node.

minimum void is selected and burst list of selected channel is updated. The burst is simply dropped if no channel is found.

Statistics in channel scheduler are wavelength utilization, burst blocking probability, calculation of offered and carried load on output link, mean number of customers and mean waiting time in scheduler queue for edge or joint node scheduling. Scheduler module is also responsible for link status update messages to router module in case adaptive routing is selected by the user. Moreover, for all scheduling algorithms implemented, the user can select either full wavelength conversion or no wavelength conversion available in scheduler module.

3.1.6 Signaling Scheme

IBKSim uses one way signaling scheme, just enough time (JET) proposed in [Qiao99]. The one way reservation is more efficient than the two way signaling, with the objective of reducing end-to-end delay of the bursts. The offset time is initially set by routing module according to the number of hops the burst has to traverse, but can be modified by scheduler to enforce QoS differentiation.

3.1.7 Routing

Each node in the network has one routing module, derived from routing layer in IBKSim. Burst arrival in every node goes to the routing module, which decides the next node of the incoming burst. A number of routing algorithms, path calculation and selection techniques are implemented in IBKSim, such as, Dijkstra algorithm, k shortest paths, k shortest loopless paths and k shortest link disjoint paths. OBS routing is discussed in detail in Chapter 5. At the time of start of simulation, or network startup, paths are calculated, and routing table is updated. OBS routing module has four functions:

Path Calculation

First, the routing module computes the paths between all source destination pairs according to input algorithm. Paths can later be computed periodically or on every connection request. Path calculation during simulation depends upon the selection of routing scheme.

Status Updates

Routing module recieves periodic status updates from connected schedulers. The contents of the status messgage depend on the routing strategy adopted in the input configuration file. The routing tables are updated by the status message information, such as, link load, utilization, drop rate, etc. The routing module later utilizes this information to recompute the paths or in path selection for assembled bursts.

Path Selection

When more than one paths are calculated in advance for every source destination pair, for example, k shortest link disjoint paths, the best one is selected for burst transmission. This selection is based on the criteria defined in the routing scheme selected for simulation. Path selection is usually applied to locally assembled bursts at the edge part of OBS network.

Burst Forwarding

In the core part of the OBS network, burst is forwarded to the next node instead of path calculation or selection. In case deflection routing is applicable, then core nodes can also make path selection. Deflection routing is not currently implemented in IBKSim.

3.2 Statistics

The statistics are collected during simulation run and are logged in the output file at given time intervals with the time stamp or once at the end of the simulation run. The simulator also supports terminal output if selected by the user, sometimes required to debug the simulation run.

It is important to specify that the logging function is called only for those components for which user has specified, and therefore, speeding up the simulation by reducing unnecessary file output. The output logging file is an XML output file, which is a generic format and very useful for post simulation processing and analysis both for the user or any program using this output.

To determine the accuracy of the output data, IBKSim divides the total simulation time/run into phases input by the user. The user also specifies the number of warm-up phases. For each of these phases independent measurements are taken, from this data, averages with corresponding confidence intervals are calculated using a Student's t-distribution. The user can select from 90%, 95% or 99% confidence interval available in the simulator. Note that between the phases no new simulation run is started, only the collected statistics over the last interval are reset. By doing this, there is no complete independence between the simulation phases, it is however, common simulation practice [Schuringa04].

3.3 Simulator Validation

Validation is the key issue to entrust the use of the results given by any simulator [Rodrigues04]. To validate the output of IBKSim, the behavior of a single link under the transit traffic is analyzed using simulations.

The local burst traffic is generated typically by an aggregate output process from a number of burst assembly queues. The output process of each individual assembly queue is renewal [Hayat11a]. It has been proven by [Lewis72] that if many renewal processes are aggregated, the combined process is exponentially distributed if aggregated process approaches to infinity in number. This assumption is valid practically in the case of an OBS assembler. It can be observed from the simulated results shown in Figure 3.8, where the pdf of burst inter-departure times from an aggregate of assembly queues have been plotted. Assembly queues are varying from one to sixteen. It can be observed that even combining output of a small number of queues for example four only, leads to an output process having a negative exponential distribution. The same assumption is true for the transit traffic by extending the comment for a big network, where the traffic is being generated independently in all nodes over the whole network.

For the behavior of an output link under transit traffic only, a simulation study has been performed. Negative exponentially distributed burst lengths and inter-arrival times have been assumed. The channel scheduling algorithm used is LAUC-VF described in Section 3.1.5 with full wavelength conversion capability available.

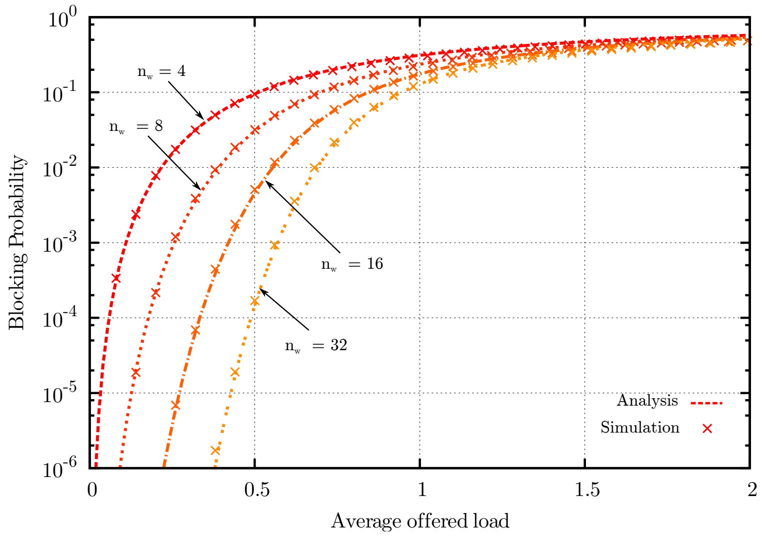


Figure 3.7: Blocking probability of a single output link in an OBS network versus the offered link load: comparison with Erlang-B loss formula.

Figure 3.7 shows the burst blocking probability as a function of offered load for the OBS output link. It can be observed that simulation output is in a close rage of those from Erlang-B loss formula for different number of wavelengths. These results validate the output of the IBKSim simulator, more specically the output of the scheduling algorithm. Traffic generation and burst assembly output also prove to be accurate.

3.4 Results

Simulations were performed on one output WDM link of OBS network, each link 16 wavelengths (specified if changed), with 10 Gbits[1] capacity. Burst are assembled using hybrid burst assembly described in Section 3.1.4.

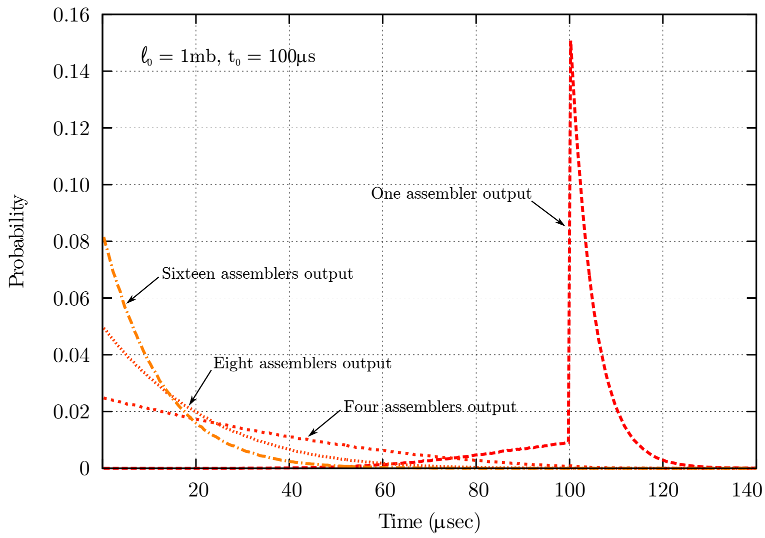


Figure 3.8: Pdf of burst inter-departure times for different number of assembly queues using hybrid burst assembly, load = 0.5.

Figure 3.8 shows pdf of burst inter-departure time from one to sixteen assemblers. It can be observed that combining multiple number of queues leads to an output process having a negative exponential distribution. The curve showing one assembler output has a vertical line at time 100, this is because of the time threshold of 100μsec.

Figure 3.9 plots pdf of burst inter-departure time with fixed length threshold and varying time threshold. The curve with minimum time threshold has more bursts assemblies with time out of time threshold.

[1] 1 Gbits in this simulation is equal to 1×10^9, i.e., 1 Gbits = 1000 Mbits exactly.

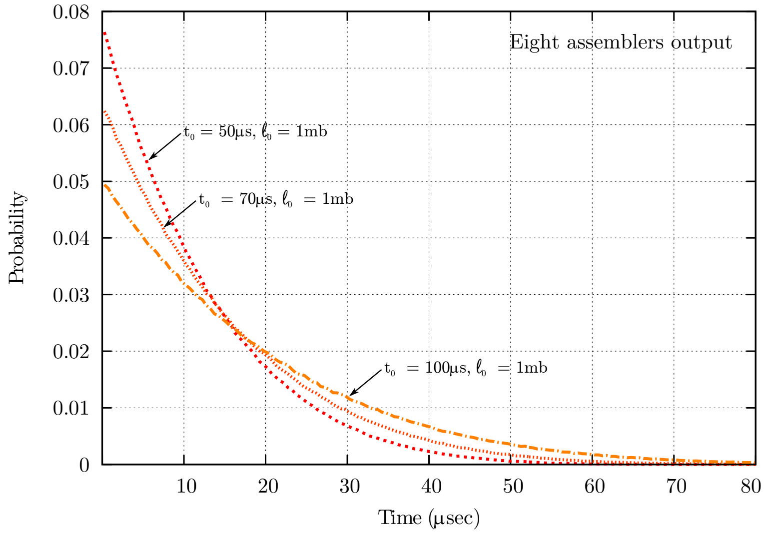


Figure 3.9: Pdf of burst inter-departure times for different time and size thresholds with eight assemblers output using hybrid burst assembly, load = 0.5.

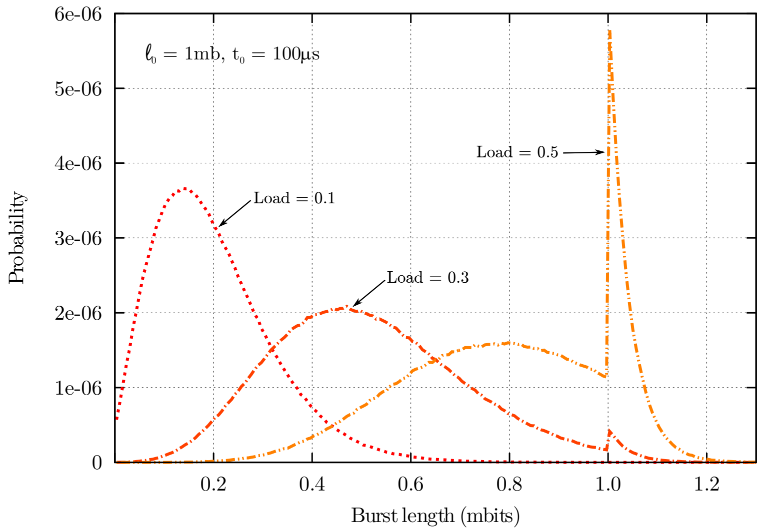


Figure 3.10: Pdf of burst length for different load scenarios using hybrid burst assembly.

Figure 3.10 draws pdf of burst length with different load scenarios. At load value 0.1, all of the bursts are assembled with time based bursty assembly, whereas, at load 0.5, bursts are assembled both with time and size based

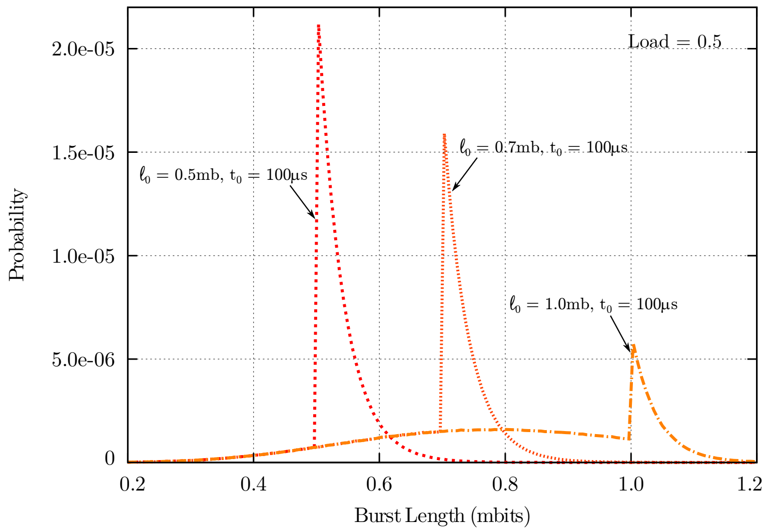


Figure 3.11: Pdf of burst length for different time and size thresholds using hybrid burst assembly.

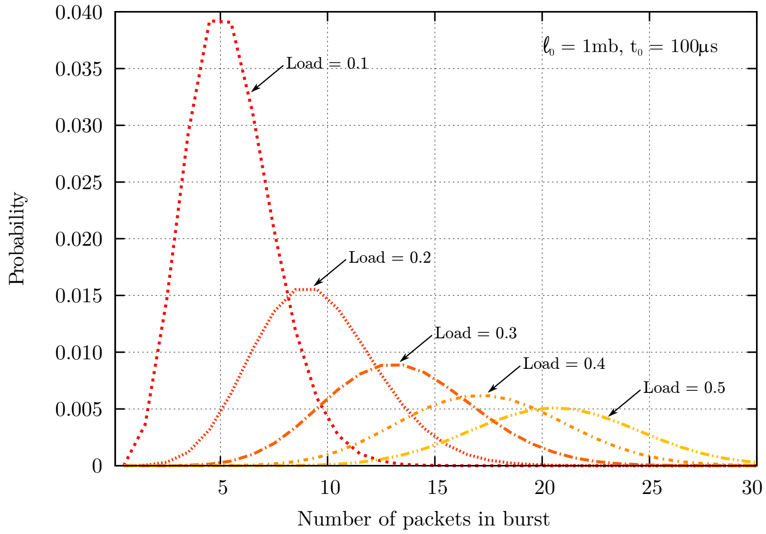


Figure 3.12: Pdf of number of packets in a burst for different load scenarios using hybrid burst assembly.

burst assembly, depending upon which threshold is met first. Figure 3.11 demonstrates pdf of burst length at fixed load and different size thresholds.

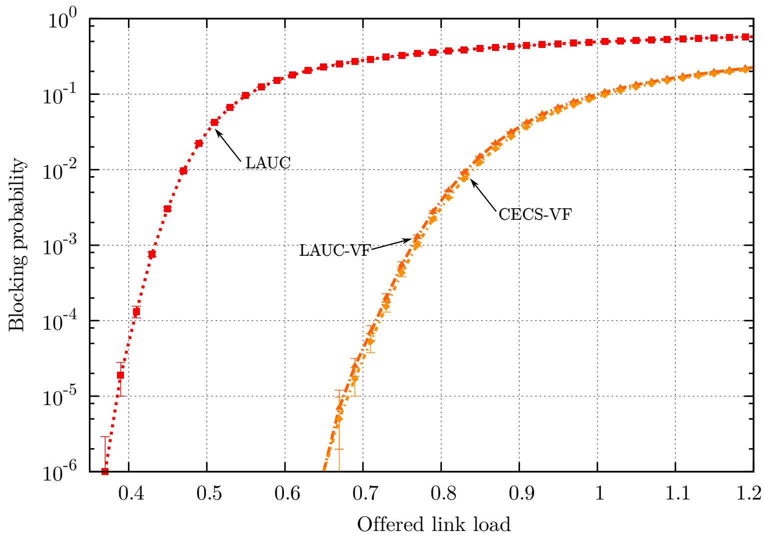


Figure 3.13: Burst blocking probability at one output link for different scheduling algorithms.

Figure 3.12 shows pdf of number of packets in a burst at different load values. The higher the load, more packets are there in the burst. Figure 3.13 shows burst blocking probability with different scheduling algorithms implemented. Results show that use of void filling scheduling algorithm improves burst blocking probability.

3.5 Summary

In this chapter we have discribed the simulation environment developed to assist the analysis of OBS networks. The components and their connectivity within a node is discussed. High level view of burst assembly and scheduling is presented.

The output of the simulator is validated with comparison of blocking probability with Erlang-B formula. Traffic characteristics of single output link are also presented in the results section.

Chapter 4

Edge-Core Joint Node

Since the paradigm of optical burst switching (OBS) was introduced over a decade ago, its performance evaluation received a considerable attention which does not seem to be weaken. Before that the full optical packet switching is technically realized, OBS will remain the most reasonable option to facilitate the statistical multiplexing over WDM. Mechanisms and functionalities considered in OBS for packet traffic burstification and for burst contention resolution provide novel and interesting scenarios for stochastic modelling. There exist already a vast literature devoted to burst assembly, optical delay lines, wavelength conversion or burst dropping strategies. Due to the system complexity, these mechanisms are usually investigated separately to determine their individual impact on burst delay and loss characteristics. Their joint performance and full inter-action are usually difficult to evaluate in an analytical way and have to be examined by simulations.

In a prevalent number of these studies, it is taken for granted that an OBS network is strictly divided into the edge and the core part. This means that the network consists of the nodes only assembling packet traffic into bursts (edge nodes) and the nodes only switching the burst traffic along the transmission path (core nodes), see Figure 4.1a. However, this assumption cannot be said valid for practical future deployments in dynamically reconfigurable networks with mesh topology. In these networks, not only a few but probably a majority of the nodes will have to combine both edge and core functionalities to provide flexible operation, as shown in Figure 4.1b. Such nodes are termed herein edge-core joint OBS nodes.

A significant challenge in the design of joint nodes is that the local traffic must be multiplexed on output wavelengths channels with the transit traffic cutting through the node, and a mechanism is needed to control the channel sharing. Otherwise, a high local load may cause high losses of the external traffic, and vice versa, too intensive transit traffic may greatly delay the transmission of locally assembled bursts. Both phenomena degrade performance of OBS. However, the increase of losses for the transit traffic is expected to have more adverse effects. Bursts that are lost not only waste the reserved path bandwidth but also invoke the retransmission and reordering delays in higher network layers. These effects are well-documented in numerous studies on Transmission Control Protocol over OBS, [Callegati06, Bikram10, Gunreben11]. On the contrary, waiting in the transmission buffer, even when excessive, is easier to monitor and does not propagate out of OBS layer to such an extent, thus assuring more stability to particular end-to-end packet flows. Therefore, when the contention for the channels occurs, transit traffic should be prioritized.

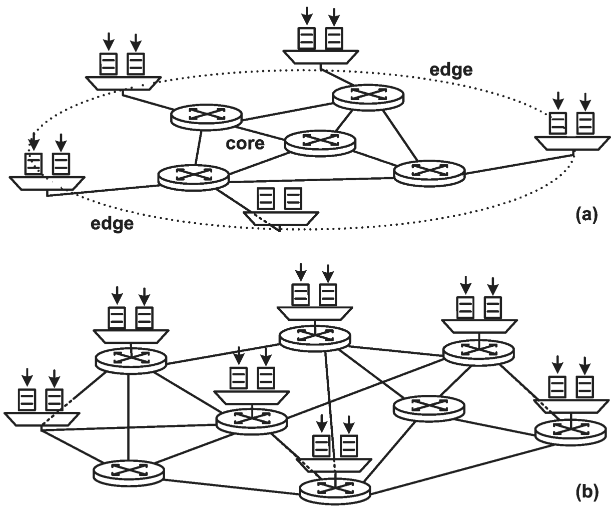


Figure 4.1: OBS network with (a) separation of edge and core parts (b) combined node

To the best of our knowledge, the issue of multiplexing of local and transit traffic, and effect of both type of traffic in OBS has not been a subject of analytical treatments. In several simulation studies, this fact was indeed assumed, however its influence on the final results was not discussed [Lee06, Barradas09, Barradas10, Yuan11]. From the theoretical viewpoint, the combined node can be modelled as a mixed loss-delay queueing system.

This class of queues has the common feature that a pool of servers handles two streams of arrivals called delay and non-delay customers. Upon finding all the servers busy, the non-delay customers are immediately lost and the delay ones are queued to wait for the service. These systems, in different variants, have found application in modelling networks nodes with integrated data and voice transmission.

Remaining in this chapter we present the architecture of edge-core joint node (ECJN) in detail, scheduling issues in the joint node, effect of high transit traffic on the delay of local traffic and effect of high local load on blocking of transit traffic. At the end a restricted intermediate buffering (RIB) in OBS networks is proposed to reduce burst loss rate which also improves fairness among drop rate of bursts with different hop count.

4.1 Edge-Core Joint Node (ECJN) Architecture

Edge-core joint node of OBS network is a combination of both edge and core nodes. It can assemble data bursts (DB) with the edge node functions and forward the transit bursts to the next node with the core node function. Figure 4.2 shows the complete architecture of ECJN. Packets arriving from different traffic sources are classified according to their traffic class and destination adress and distributed into corresponding assembly queues. The bursts are assembled in edge node, according to size, time or hybrid based burst assembly mechanisms. The locally assembled bursts are then forwarded to burst transmission queues (BTQs), where they are buffered electronically and wait for the availability of the output channel. BTQs are explained in detail in Section 4.1.4. When the scheduling module finds output channel, it reserves the channel and forwards burst control packet (BCP) to the destination node, and forwards the burst from BTQ to the scheduler buffer, where the bursts wait electronically until their transmission time is reached.

On the other hand, while performing the functions of the core node, the transit traffic (BCP) arrives at routing module, if the destination of arriving bursts is current node, the bursts are forwarded to the burst disassembly module, which disassembles the bursts into IP packets and forwards IP packets to their destinations. If transit bursts need to be forwarded to the next nodes, information is sent to the scheduling module, which looks for the availability to reserve output wavelength channel. If channel is found after wavelength conversion (if required), the wavelength is reserved for incoming bursts, otherwise the burst is dropped.

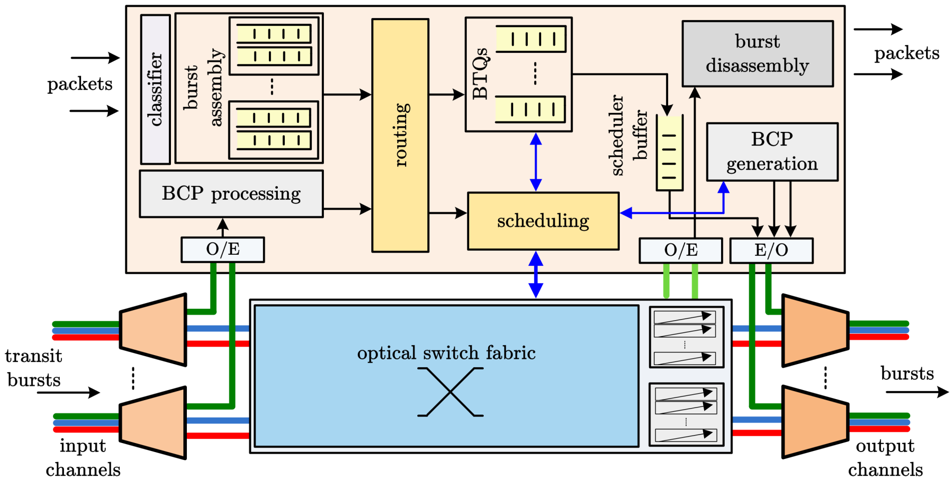


Figure 4.2: Architecture of Edge-Core Joint OBS Node

Components of edge-core joint node are discussed in detail

4.1.1 Classifier

The classifier recieves the incoming packets from different packet sources, retrieves the packet header to extract destination, class of service and quality of service information and forwards the packet to the relevant assembler queue.

4.1.2 Burst Assembly

The burst assembly in a joint node is similar to the burst assembly in an edge node. After sorting of packets into assembler queues by the classifier, the burst assembly process starts from the arrival of first packet in each queue. The number of assembler queues depends upon the number of destination nodes in the network, and class of service. There exists one queue for each class of service for every destination node. The bursts are assembled according to size, time or hybrid based burst assembly technique.

4.1.3 Routing

The routing module in a joint node has to perform both edge and core node functions. As an edge node, when a new burst arrives, the routing module selects the most feasible path for the burst's destination, from the existing pre-calculated paths, or calcualtes a new path on each burst arrival. The decision of selection or calculation depends upon the adopted routing strategy. After the routing decision, the burst is forwarded to the scheduler module of the selected path's output link. In case of arrival of control packet of transit burst, while performing the function of core node's routing module, next hop of the burst is retrieved from the routing table and the control packet is forwarded to the scheduler module of next hop's output link. In case that the destination of transit burst is current node, the incoming burst is forwarded to the burst disassembly module, where the packets in the burst are disassembled and forwarded to the respective packet destinations.

4.1.4 Burst Transmission Queues (BTQs)

The locally assembled bursts dispatched from routing module to the scheduler module arrive into the burst transmission queues (BTQs) while the scheduler module searches for output wavelength for scheduling. Ideally there should be an infinite buffer availability for the delay traffic, but a higher load of transit traffic results in the loss of the locally assembled bursts in case the buffer is full. Whereas, the transit bursts have no buffering facility available. BTQs are further explained in channel scheduling in Section 4.2.

4.1.5 Scheduling

Scheduling in a joint node is a combination of scheduling in an edge node and the core node. The core node operates in all optical domain, at the arrival of a control packet, the scheduler module searches for the time slot to fit the burst on the output wavelength channel. If the wavelength is found (possibly with the need of wavelength conversion), the output wavelength is reserved for arriving burst, starting from arrival time of the burst. Otherwise, if channel is not found, the burst is dropped due to the unavailablity of optical buffers. At most, fiber delay line (FDL) components may be used but FDLs do not operate similar to the electronic buffers.

As discussed in Section 4.1.4, the local burst data remains in the electronic domain during scheduling, therefore, rather than dropping the bursts, it is possible to buffer them until channel is available. The scheduler module searches for availability of channel, when it is found, the burst from the front of the queue is forwarded to the scheduler buffer, and channel is reserved for the considered burst after its offset time. Channel scheduling for the joint node is explained in detail in Section 4.2.

4.1.6 Buffer

In the joint node, there are three locations which require electronic buffers.

1. Assembly queues
2. Burst transmission queues
3. Scheduler buffer

The assembly queues themselves are buffers, packets are delayed until burst completion according to the burst assembly algorithm implemented. Burst transmission queues are the queues for local bursts, and scheduler buffer is the queue where local bursts are delayed from their control packet generation to the actual transmission.

4.1.7 Control Packet Processing and Generation

The control packet for the transit bursts is received after opto-electrical conversion, and forwarded to the routing module. Control packet generation module generates the control packet for the local assembled burst after the wavelength reservation by scheduling unit. This control packet contains necessary information of preceding burst, for example, burst arrival time, burst length, burst destination, selected path index and class of service. The primary function of the control packet is to reserve the resources in advance for the incoming burst.

4.2 Channel Scheduling in Edge-Core Joint Node

Channel scheduling is used to allocate the suitable output wavelength channels to the incoming bursts. Scheduling in an edge node is different than scheduling in the core node. The edge node is capable to buffer the bursts in the electronic domain if there is no output wavelength channel available. Whereas, the core node operates in the optical domain, and if scheduling algorithm fails to find any output wavelength channel for the burst, the burst is simply dropped. The fiber delay lines are proposed to be used as buffers in the core node, but they work differently than queueing buffers.

The channel scheduling in a joint node is different than both edge and core node. In a joint node, locally assembled traffic has to compete with the transit traffic for the output wavelengths. Similarly, the transit bursts coming from previous nodes have to compete for output resources with the local traffic. The transit bursts are dealt as loss traffic and therefore, the burst is dropped if the scheduling algorithm remains unable to find a time slot for the burst to fit on the output channel. Figure 4.3 describes the queueing model of input traffic in a joint node. The arrivals of the transit traffic go directly to the wavelength channels, whereas, the local bursts have the buffering facility.

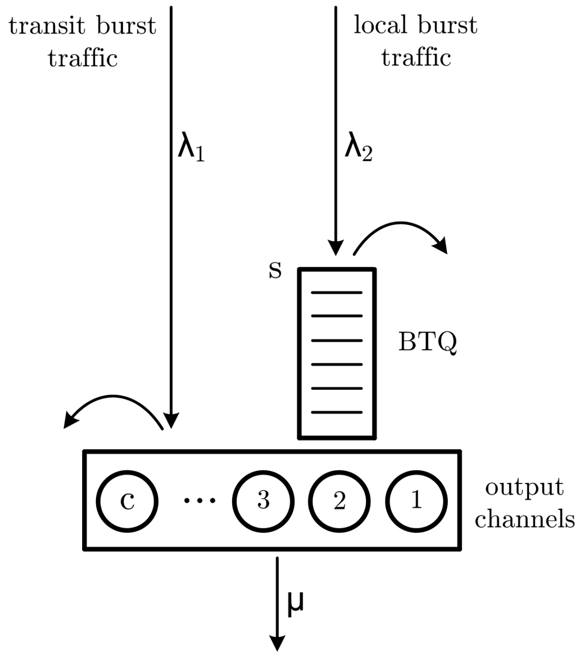


Figure 4.3: Queueing model for a single output port of the OBS edge-core joint node

A new burst scheduling algorithm, composite edge core scheduling with void filling (CECS-VF) is proposed to schedule both local and trasit bursts. This algorithm is based on latest available unused channel with void filling (LAUC-VF) proposed in [Xiong00], it takes into account the burst transmission queues where local assembled bursts are waiting and also schedules the transit bursts.

4.2.1 Composite Edge/Core Scheduling (CECS-VF)

The CECS-VF performs a composite scheduling of both local and transit bursts. This scheduling algorithm is different from the general edge sched-

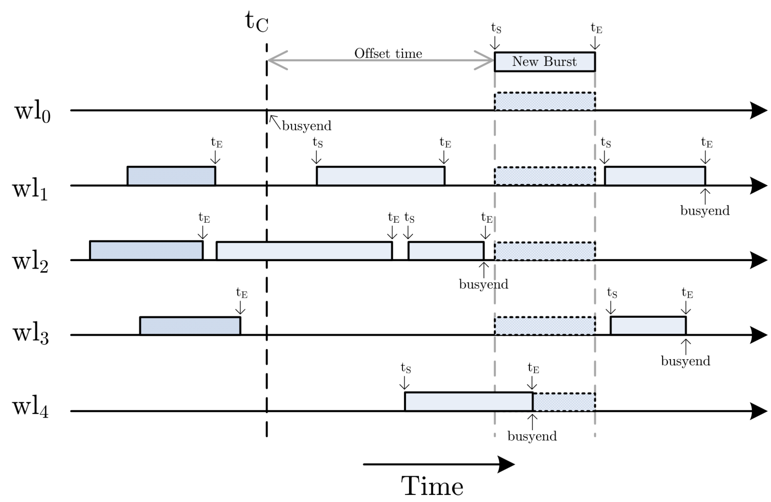


Figure 4.4: Scheduling example of new burst for a signle output port of OBS node

uler in the sense that it searches and reserves wavelength when there is the time to transmit burst control packet (BCP). Earliest possible minimum void with void filling (EPMV-VF) proposed in [Li07], is an edge scheduler algorithm which reserves wavelengths in advance prior to BCP transmission time. And therefore, it keeps track of burst control packet time t_{cp}, and transmits BCP when t_{cp} occurs. This practice if implemented in the composite scheduling technique will give priority to the local traffic, and increase the loss rate of transit traffic. The offset time of local burst is already calculated in the router module according to the number of hops this burst have to traverse, therefore, when the burst arrives, the scheduler finds for the available time slot for this burst to fit on the wavelength channel. The CECS-VF algorithm is given in Algorithm 4.2.1 and Algorithm 4.2.2.

In CECS-VF, whenever there is an arrival, either local burst or control packet for transit burst, the BurstARrival method is called. The burst received from the local router module is inserted into BTQ, then the reference of the burst at front of the queue is retrieved. The scheduler module searches for the availability of output channel after the offset time of the burst from burst start time $t_S[B_i]$ to burst end time $t_E[B_i]$ using LAUC-VF algorithm. As shown in the Figure 4.4, channel is searched for the time slot of burst size in time after current time plus pre-calculated offset time of the burst. Three possibilities can occur, there can be no burst on the channel, burst can be scheduled in the void and the contention may occur. In case of contention, the channel is supposed to be unavailable. If no channel is found for this time slot, the burst remains into BTQ. If a channel is found, it is reserved, BCP for the burst B_i is transmitted, and the burst is moved from BTQ to the scheduler buffer and statistics are updated. One more event is scheduled at the end time of scheduled burst $t_E[B_i]$, i.e., ChannelAvailable. This event tells the scheduler that there is a void available at this moment and the output wavelength channel is available for scheduling. When scheduling the local burst, all output wavelength channels are traversed for availability and selected according to LAUC-VF.

Algorithm 4.2.1: BurstArrival(B_i)

$$
\begin{array}{l}
\textbf{if } \text{LocalBurst} \\
\quad \textbf{then } \begin{cases}
\text{Insert burst into BTQ} \\
B_i \leftarrow \text{Front burst of BTQ} \\
\text{Find channel using LAUC-VF} \\
\textbf{if } \text{Channel found} \\
\quad \textbf{then } \begin{cases}
\text{Reserve channel} \\
\text{Remove } B_i \text{ from BTQ} \\
B_i \text{ forwarded to Scheduler Buffer} \\
\text{CALL ChannelAvailable() at } t_E[B_i] \\
\text{Update statistics}
\end{cases} \\
\quad \textbf{else } \begin{cases} B_i \text{ stays in BTQ} \end{cases}
\end{cases} \\
\textbf{else if } \text{TransitBurst} \\
\quad \textbf{then } \begin{cases}
\text{Find channel using LAUC-VF} \\
\textbf{if } \text{Channel found} \\
\quad \textbf{then } \begin{cases}
\text{Reserve channel} \\
\text{CALL ChannelAvailable() at } t_E[B_i] \\
\text{Update statistics}
\end{cases} \\
\quad \textbf{else } \begin{cases}
\text{Drop the burst } B_i \\
\text{Update statistics}
\end{cases}
\end{cases}
\end{array}
$$

In case of arrival of the control packet of a transit burst, the scheduler module searches for availability of the channel for the arriving burst. If no wavelength conversion is available, the scheduler finds space on one wavelength, whereas, if switch is capable of full wavelength conversion, all wavelength channels are traversed, if no space is found on incoming burst's wavelength. If no channel is available, the burst is simply dropped. And if channel is found, it is reserved, ChannelAvailable event is scheduled at $t_E[B_i]$ and statistics are updated for both cases.

Algorithm 4.2.2: ChannelAvailable( )

---

if BTQ is not empty
then
- $B_i \leftarrow$ Front burst of BTQ
- Find channel using LAUC-VF
- **if** Channel found
  - **then**
    - Reserve channel
    - Pop B_i from BTQ
    - B_i forwarded to Scheduler Buffer
    - CALL ChannelAvailable() at $t_E[B_i]$
    - Update statistics
  - **else**
    - B_i stays in BTQ

The ChannelAvailable is called whenever there is availability of void on the output wavelength channel. If there are bursts in BTQ, this module performs same functions as the scheduling of the local burst is performed.

4.3 Performance Results

In this section we analyze the behavior of ECJN in terms of mixing of local and transit traffic using simulation environment described in Chapter 3. It is worth mentioning that higher rate of transit traffic causes increase in waiting time of local bursts in the burst transmission queues (BTQs), and vice versa, higher rate of local traffic causes higher burst loss of transit traffic.

Simulation Setup

Simulations of edge-core joint node are performed on a 5-node sample topology presented in Figure 4.5. In this topology, node $1-2-3$ are ingress edge nodes and node 4 is a joint node. All of the four nodes generate bursts which are destined to egress node 5. The local bursts assembled at joint node 4 compete for output wavelengths with transit bursts arriving from previous nodes.

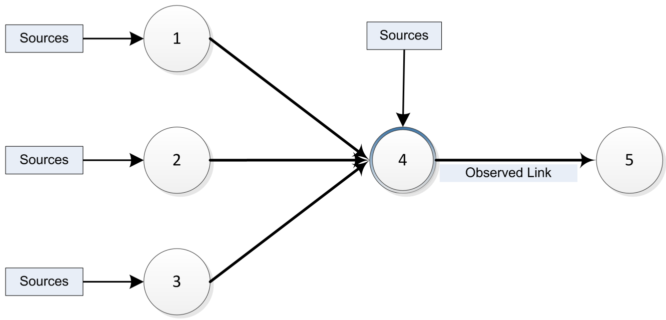


Figure 4.5: 5-node topology, with node 4 as a joint node.

Packet arrivals to the nodes are Poisson with exponentially generated mean packet length of 40 $kbits$. Bursts are assembled using hybrid based burst assembly mechanism with time threshold of 100 μs and size threshold of 1 $Mbits$[1]. CECS-VF explained in Section 4.2 is used for wavelength scheduling. Burst transmission queues (BTQs) can buffer upto 1000 bursts.

We assume the following:

- The number of wavelengths on each link is 4 with 10 $Gbit/s$ capacity.
- Full wavelength conversion is available at node 4.
- No re-attempt is performed when a connection is blocked.

[1] 1 $Mbits$ is equal to 1×10^6 bits

Calculation of offered network load and blocking probabilities are according to the OBS network modelling explained in Section 5.5. Total simulation time is divided into 20 intervals plus a warm-up or transient period and 95% confidence interval is used to determine the accuracy of the output. 0.1 million bursts are generated for each simulation interval.

Results

Figure 4.6 plots the waiting time of local bursts as a function of offered transit load while the local load is fixed at 0.2, 0.3 and 0.5. We can observe that with the increase in transit load, the waiting time of the local bursts increases. Similar behavior is shown in Figure 4.7, which plots the BTQ size for local bursts. Figure 4.8 draws the mean blocking probability of transit bursts as a function of local load. It can be observed that with the increase in local traffic load, transit traffic is affected badly in terms of burst blocking probability.

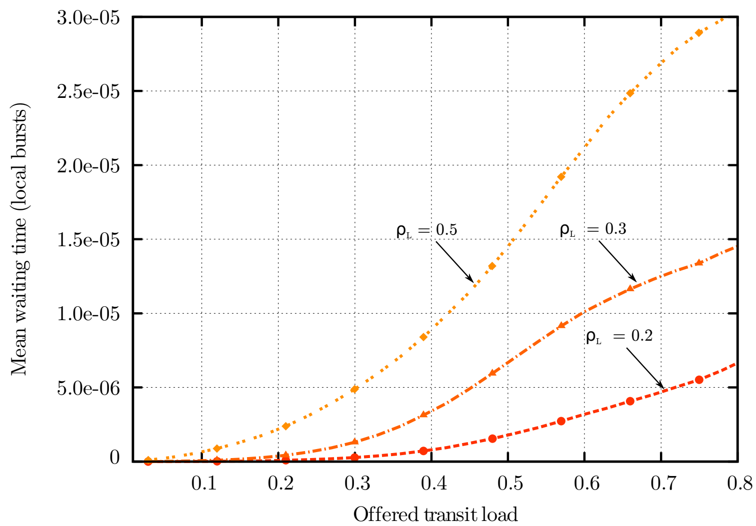


Figure 4.6: Waiting time of local bursts as a function of offered transit load.

The results show that if both local and transit traffic has equal priority, both will effect the traffic of other. The usage of wavelengths can be restricted for both local and transit traffic to achieve a bargain between mean waiting time of local traffic and the loss rate of transit traffic. Mixing of local and transit traffic is studied in [Hayat12] to control the output channel sharing by extending the basic mixed loss-delay queueing models and are solved by Markov chain techniques.

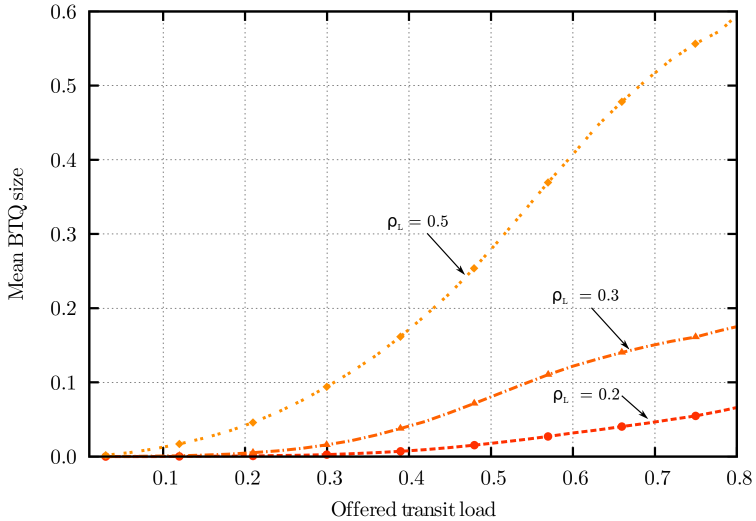


Figure 4.7: Mean BTQ size as a function of offered transit load.

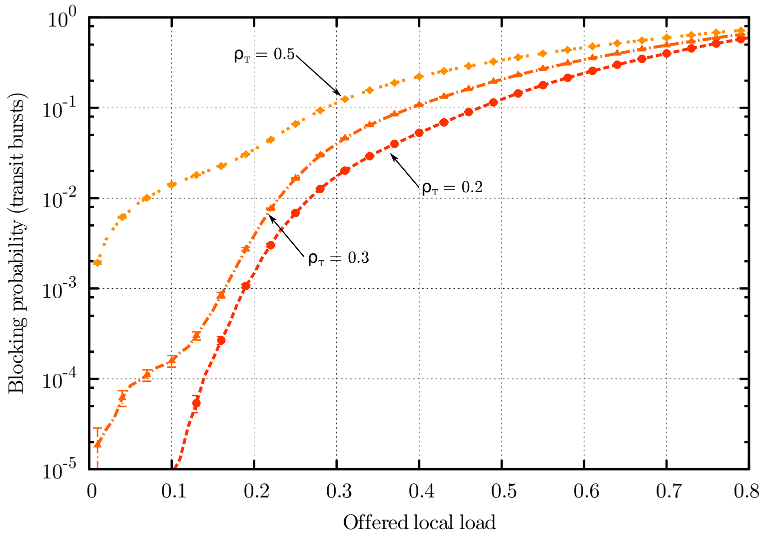


Figure 4.8: Mean blocking probability of transit bursts as a function of offered local load.

4.4 Restricted Intermediate Buffering in OBS Networks

Due to the bufferless nature of OBS networks, transit bursts in the core network are dropped if channel scheduler remains unable to find the output wavelength. The burst is the smallest data unit in OBS networks, therefore, several packets are dropped when a burst is dropped. There are many reactive mechanisms proposed for contention resolution such as wavelength conversion, use of fiber delay lines and deflection routing [Hayat11b], [Wang00]. In deflection routing the bursts are deflected to the secondary paths in case of contention on primary paths. The deflection in the network results in several side effects [Gao08] including burst transmission delay and out-of-order packet arrival at the destination. With the use of deflection routing, the offset time management requires use of fiber delay lines as well.

Next available neighbour routing algorithm (NAN) proposed in [Garcia08] is a reactive mechanism to deflect the contending burst to the nearest edge node, where the burst is converted back to electical domain, buffered and routed to the destination again at the availability of channel. NAN routing algorithm improves overall burst loss rate while a counter method is needed to avoid endless loops in the burst's route in the network.

As in the case of edge-core joint nodes, electronic buffer is available with most of the nodes in the network and the transit bursts can be buffered after conversion from optical to electrical, and reinserted into the network whenever the time slot on the output wavelength channel is available. In this way, each joint node will have one more buffer for the tarnsit bursts as shown in Figure 4.9, and this mechanism is referred to as restricted intermediate buffering (RIB) in OBS networks. The additional buffer for the transit bursts is called intermediate buffering queue (IBQ).

While performing the function of an edge node, the joint node gives priority to IBQ customers and schedules the local bursts in case when IBQ is empty to take advantage of reducing transmission time of transit bursts. If every contending burst in the core network is buffered there will not be any blocking in the core network, but this will adversly affect the transmission of local assembled bursts. Therefore, some restrictions are applied to reduce the effect of intermediate buffering on local traffic.

It is also established fact that the longer paths are exposed to higher blocking probability [Ueda05]. This problem is termed as unfairness in OBS networks i.e., the more the number of hops a burst has to traverse, the higher the probability of blocking. Moreover, when a burst is dropped on a longer path, it adversely effects the overall network performance because

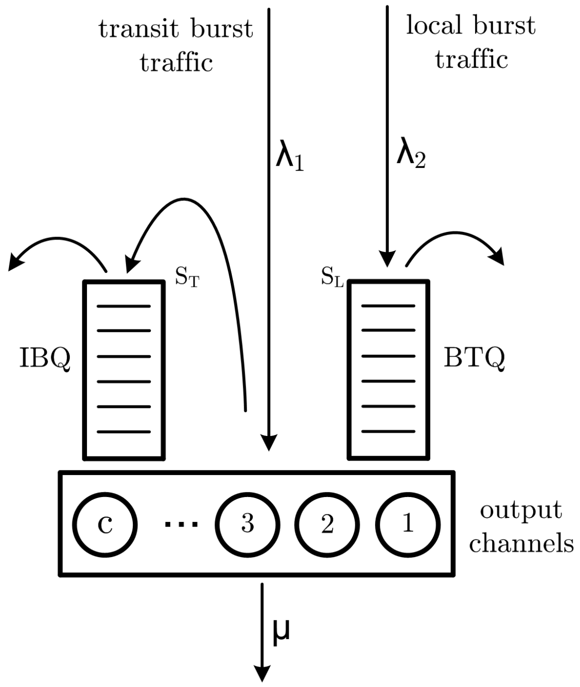


Figure 4.9: Queueing model for a single output port of the OBS edge-core joint node with restricted intermediate buffering

the dropped burst have already consumed lot of network resources. RIB deals with the OBS's inherent unfairness problem, and restricts the buffering of bursts to only those which have already consumed much resources of the network. Two intermediate buffering schemes last hop intermediate buffering (LHIB) and resource consumption based intermediate buffering (RCIB) are proposed to resolve the contention in the core part of the network and to improve fairness among bursts with different number of hop count.

4.4.1 Last Hop Intermediate Buffering (LHIB)

Last hop intermediate buffering method buffers those contending bursts which are going to be dropped while only one hop is left to their destination. If a burst has to traverse 3-hops to the destination, in JET based signaling, first two hops reserve the wavelength for incoming burst, while the third and last hop remains unable to reserve output wavelength, the burst will be dropped. Dropping at this point is very crucial because the burst will consume the resources on previous hops. To address the problem of last hop dropping, LHIB is proposed in which intermediate buffering is restricted to only last hop contending bursts.

4.4.2 Resource Consumption based Intermediate Buffering (RCIB)

In LHIB, longer path bursts may not achieve the full advantage of intermediate buffering in case the bursts face contention on the second hop. To address this problem, resource consumption based intermediate buffering (RCIB) method is proposed, which buffers those transit bursts which have traversed up to the half of network diameter. The diameter of a network is defined as the longest of all the calculated shortest paths in a network.

RCIB implies that the more the resources a burst has already consumed, the more preference should be given to this burst, because dropping of such burst is not only the loss of burst but also the loss of resources on previous nodes. Half of the network diameter is chosen to give priority to longer path bursts, and as a result, it will improve the fairness among bursts with different hop count.

4.4.3 Advantages of Intermediate Buffering

There are number of advantages gained by restricted intermediate buffering:

1. Intermediate buffering acts as a contention resolution mechanism and lowers the burst loss rate in the core network.
2. Improves fairness by prioritizing the longer path bursts.
3. Unlike fiber delay lines (FDL), the bursts can effectively be buffered (electronically) for the longer period of times.
4. Helps in reducing the retransmission of data by upper layers, and therefore attempts to lighten the load in the network.

4.4.4 Performance Results and Comparisons

The performance of proposed intermediate buffering mechanisms for contention resolution and fairness improveness is investigated using simulation environment described in Chapter 3. The simulations were performed on NSFNet based 15 node 25 links topology shown in Figure 4.10, Cost239 based 11 node 26 links topology shown in Figure 4.11, and 16 node 32 links 4x4 Torus topology shown in Figure 4.12. All of the links in given topologies are bidirectional links.

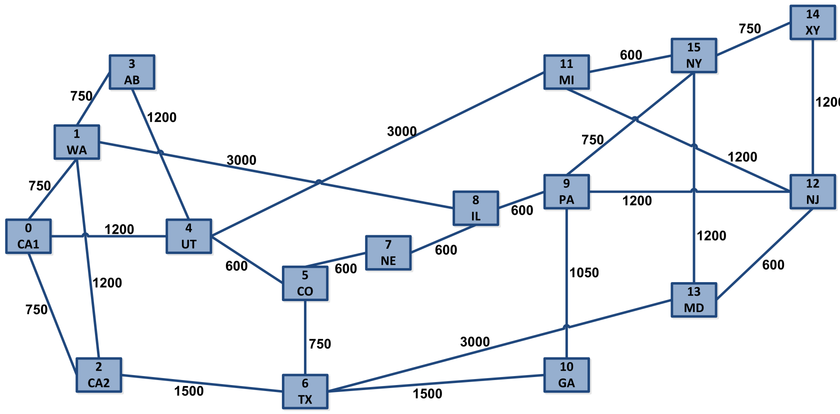


Figure 4.10: NSFNet based 15 node 25 bidirectional links topology

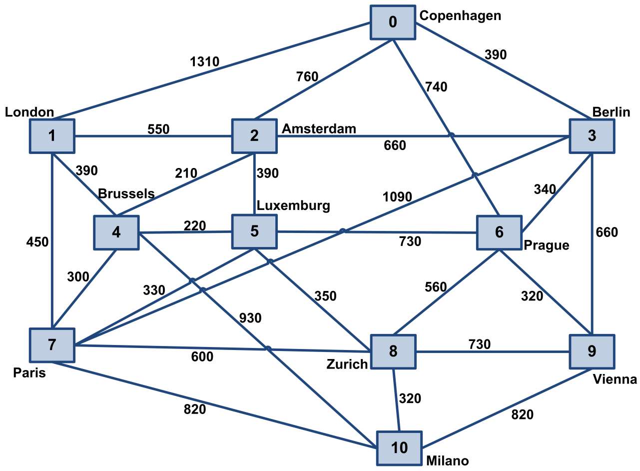


Figure 4.11: Cost239 based 11 node 26 bidirectional links topology

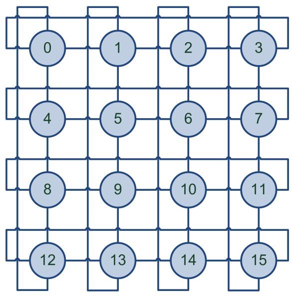


Figure 4.12: 16 node 32 bidirectional links 4x4Torus topology, link length=10 km

Simulation Parameters

All nodes of the network are joint edge-core nodes as described in Section 4.1 which are connected with packet sources, packet arrivals to the nodes are Poisson with exponentially generated mean packet length of 40 $kbits$. The bursts are assembled using hybrid-based burst assembly mechanism with time threshold of 100 μs and size threshold of 1 $Mbits$. CECS-VF explained in Section 4.2 is used for wavelength scheduling. Burst transmission queues (BTQ) and intermediate buffering queues (IBQ) can buffer upto 1000 bursts. Dijkstra algorithm [Dijkstra59] is used for the computation of shortest paths between all source destination pairs.

We assume the following:

- The number of wavelengths on each link is 16 with 10 $Gbit/s$ capacity.
- The traffic is uniformly distributed among all node pairs.
- Control packet processing time at a node is 10 μs.
- The offset time of each burst is set according to the number of hops the burst has to traverse. At each hop, offset time is reduced by control pcaket processing time i.e., 10 μs.
- The shortest path obtained in routing is defined as the path with the minimum number of hops.
- The propagation delay per kilometer is 5 μs.
- Full wavelength conversion is available at each node.
- No re-attempt is performed when a connection is blocked.

Calculation of offered network load and blocking probabilities are according to the OBS network modelling explained in Section 5.5. Total simulation time is divided into 20 intervals plus a warm-up or transient period and 95% confidence interval used to determine the accuracy of the output. 0.1 million bursts are generated for each simulation interval.

Results

Figures 4.13, 4.14 and 4.15 plot the overall network loss in terms of mean blocking probability against average offered load per link for NSFNet, Cost239 and 4x4 Torus network topologies. The blocking probability results of LHIB and RCIB are compared with no intermediate buffering. It can be realized that the LHIB performs better in terms of burst loss in all three networks. In case of Cost239 network, Figure 4.14, LHIB shows much better results and there is no blocking when load is less than 0.48. The performance of both proposed schemes depends upon network connectivity and average hop degree. Table 4.1 shows hop count for shortest path routing for all three topologies.

Network	Number of hops		Network
	Average	St. Deviation	Diameter
NSFNet	2.14	1.01	4
Cost239	1.42	0.50	3
4x4 Torus	2.00	1.00	4

Table 4.1: Number of hops in shortest paths for all three topologies.

The Cost239 network has more connectivity and average hop distance is 1.42 hops, therefore, the paths consisting of 2-hops get the advantage of intermediate buffering with LHIB scheme. Network diameter for this network is 3, but very few paths are comprised of 3-hops, therefore RCIB with half of network diameter algorithm could not perform better in terms of network burst blocking probability. On the other hand NSFNet and 4x4 Torus networks both have network diameter of 4, therefore, the bursts traversing more than two hops were able to utilize the benefits of intermediate buffering. In these networks, LHIB algorithm performs better as well.

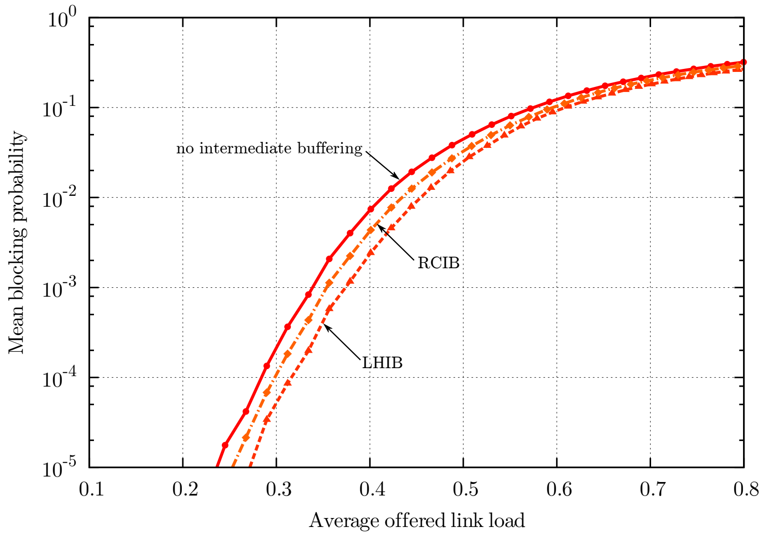


Figure 4.13: Mean blocking probability as a function of average offered link load for NSFNet network.

Figures 4.16, 4.17 and 4.18 show the mean burst transmission queue (BTQ) size as a function of average offered link load. BTQs and their function is discussed in detail in Section 4.2, as each link has one scheduler and every

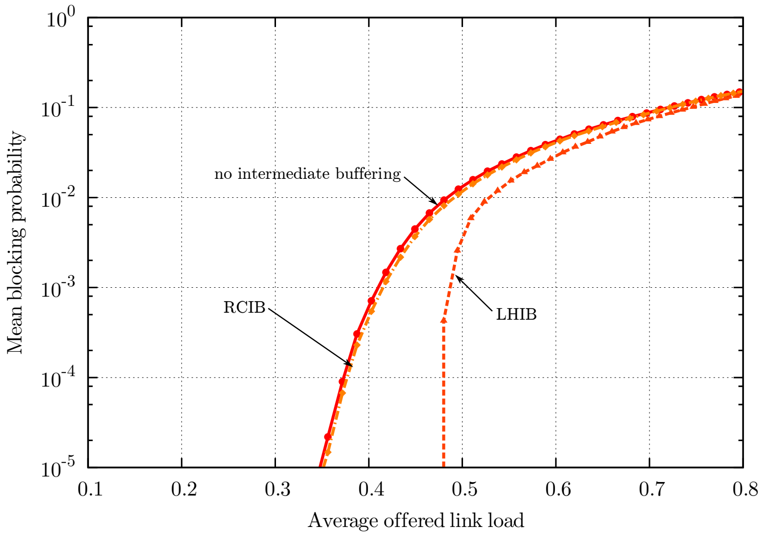


Figure 4.14: Mean blocking probability as a function of average offered link load for Cost239 network.

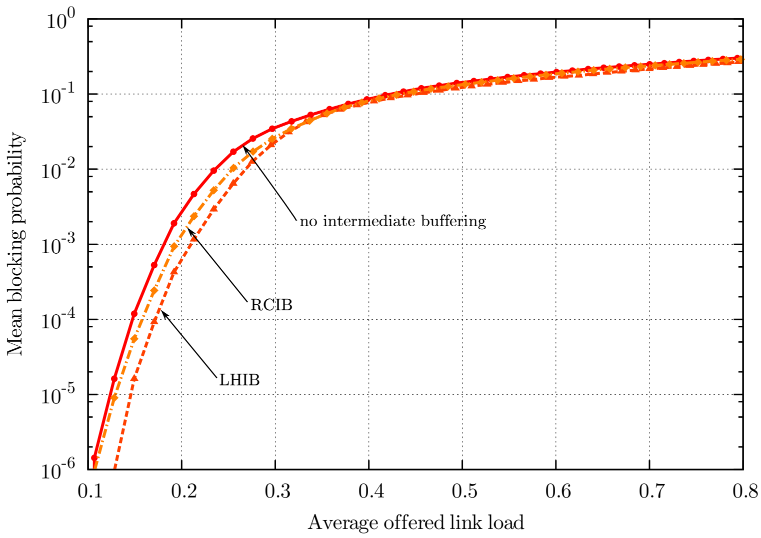


Figure 4.15: Mean blocking probability as a function of average offered link load for 4x4 Torus network.

scheduler has one BTQ in a joint edge-core node. It is important to mention that when intermediate buffering is used, priority is given to intermediate buffering queue (IBQ), BTQ is only sevred when IBQ is empty, as a result, number of customers in BTQs increase. Blocking probability is better in

LHIB algorithm and therefore, mean BTQ size for LHIB is also more than RCIB and without intermediate buffering strategies.

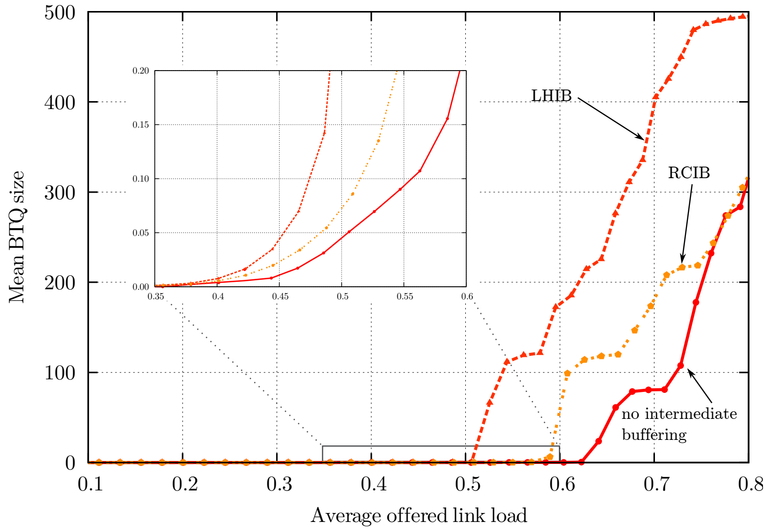


Figure 4.16: Mean BTQ size as a function of average offered link load for NSFNet network.

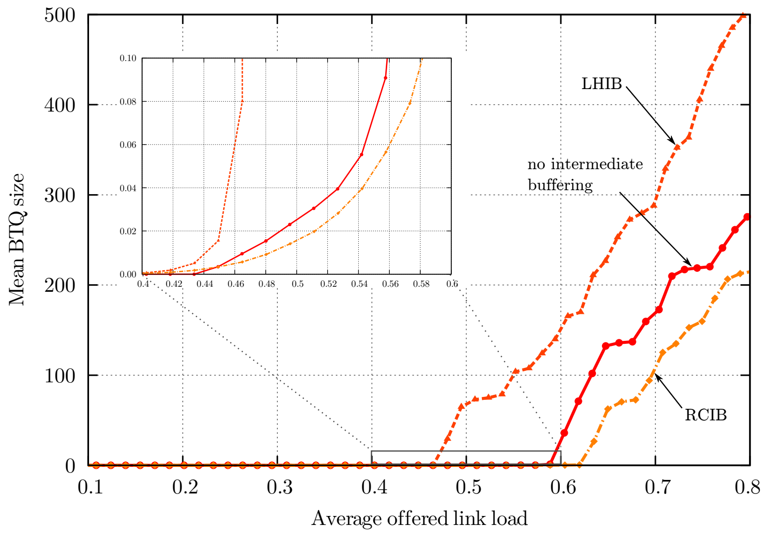


Figure 4.17: Mean BTQ size as a function of average offered link load for Cost239 network.

The jumps in queue size curve are due to the fact that at every step, a number of BTQs reach to their maximum capacity, i.e., 1000 bursts. This

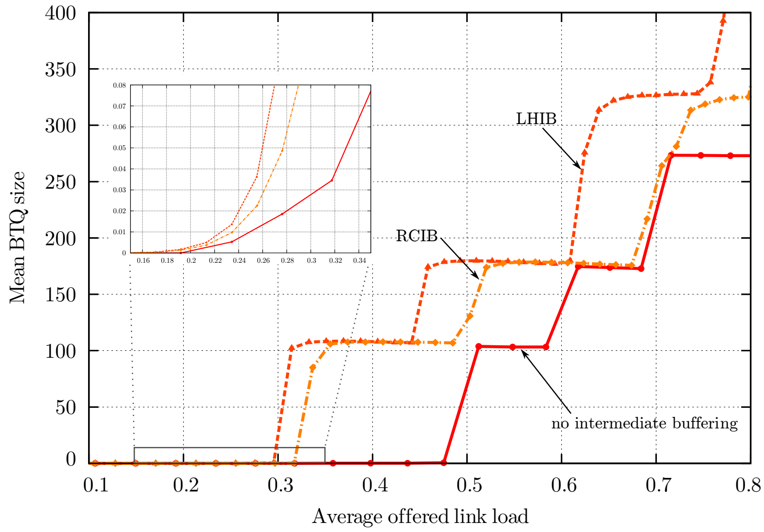


Figure 4.18: Mean BTQ size as a function of average offered link load for 4x4 Torus network.

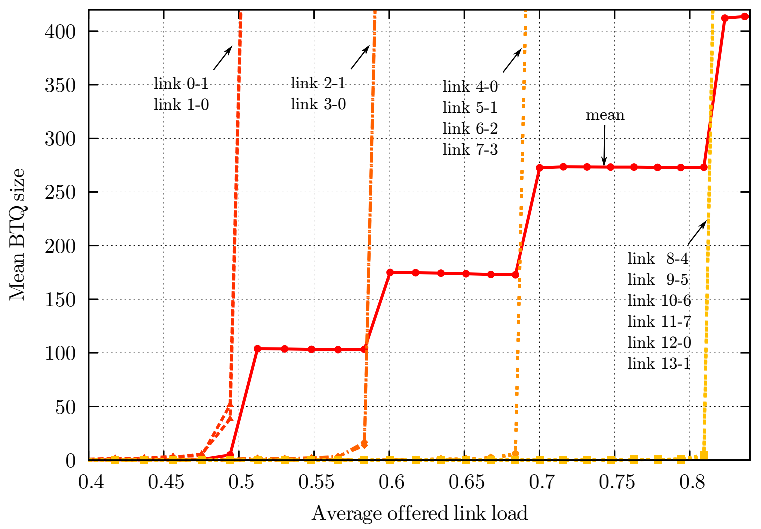


Figure 4.19: Mean BTQ size as a function of average offered link load for 4x4 Torus network with no intermediate buffering.

effect is more visible in Figure 4.18 because of the reason that this network has more symmetry in connectivity. Schedulers with similar number of usage in shortest paths get overloaded simultaneously. Figure 4.19 shows some of the BTQs which get full and steps appear in mean queue size. The more the BTQs get full, the bigger is the resultant step. Links $0 - 1$ and $1 - 0$ have same number of usage in shortest path routing, therefore the BTQs attached with these links reach to their maximum at the same point.

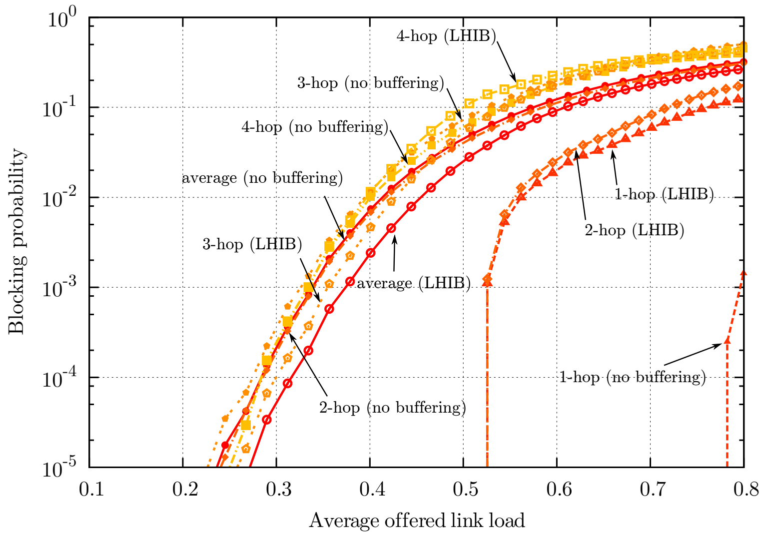


Figure 4.20: Hop wise blocking probability as a function of average offered link load in LHIB scheme for NSFNet network.

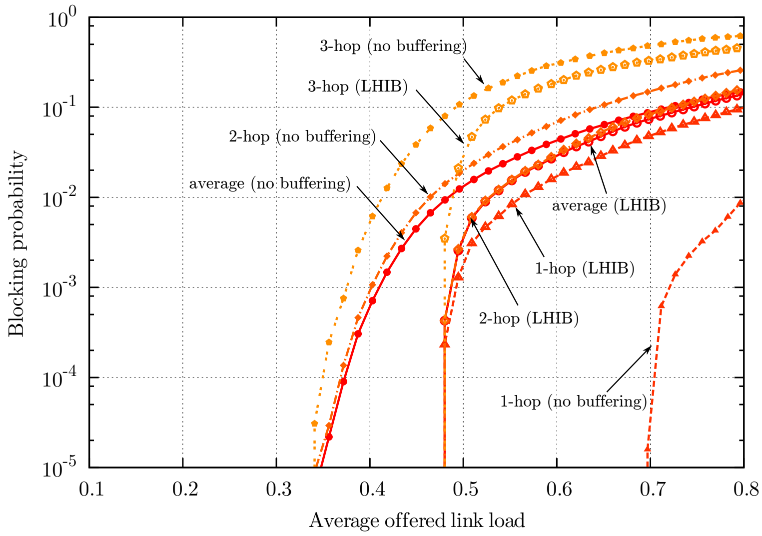


Figure 4.21: Hop wise blocking probability as a function of average offered link load in LHIB scheme for Cost239 network.

Figures 4.20, 4.21 and 4.22 demonstrates the burst loss performance based

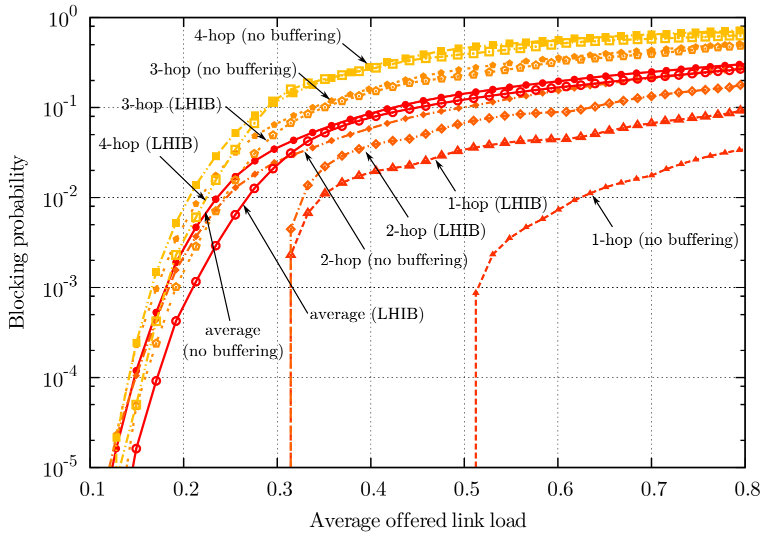


Figure 4.22: Hope wise blocking probability as a function of average offered link load in LHIB scheme for 4x4 Torus network.

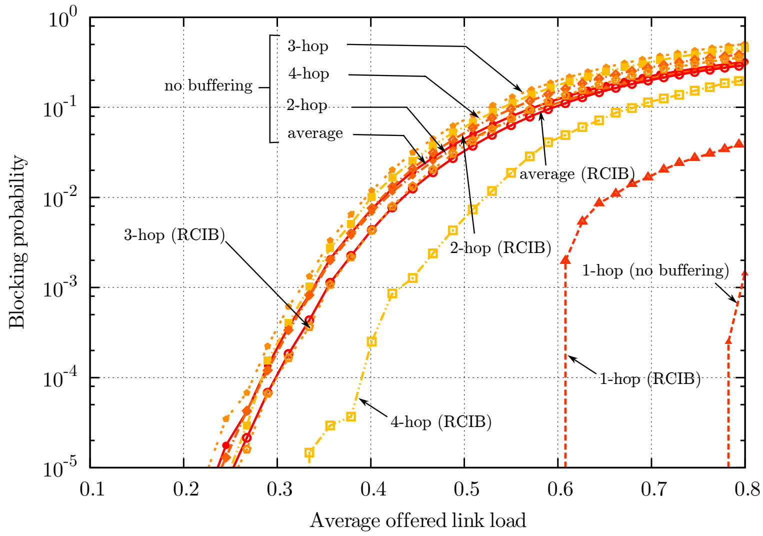


Figure 4.23: Hope wise blocking probability as a function of average offered link load in RCIB scheme for NSFNet network.

on number of hops for LHIB for all three network topologies. It can be observed from the figures that 2-hop bursts are benefited the most using last hop intermediate buffering. While 3-hop and 4-hop traffic loss is above the average curve. The LHIB scheme benefits the 2-hop traffic in the sense

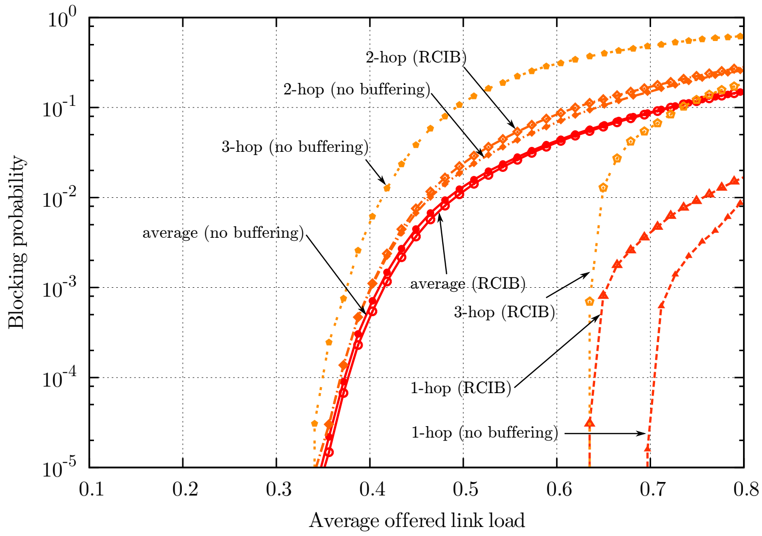


Figure 4.24: Hope wise blocking probability as a function of average offered link load in RCIB scheme for Cost239 network.

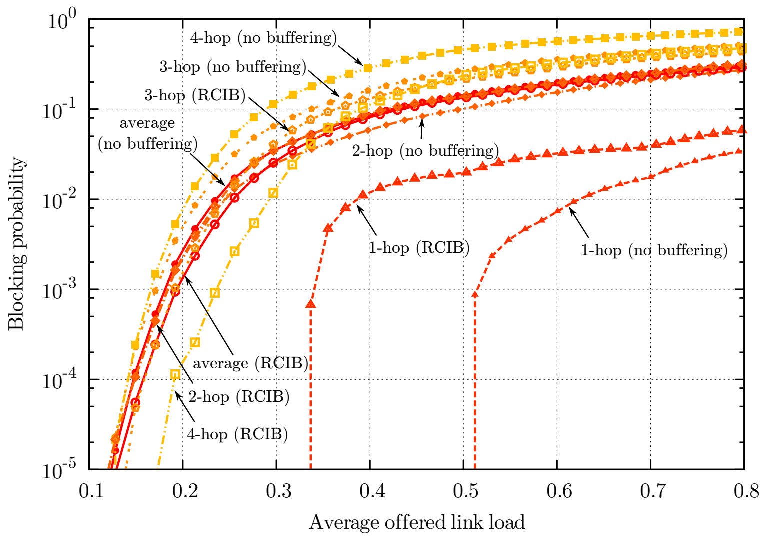


Figure 4.25: Hope wise blocking probability as a function of average offered link load in RCIB scheme for 4x4 Torus network.

that the bursts get the facility of buffering on both hops, and rare chance of drop is left. As a result, the 2-hop bursts are not dropped until the load is slightly above 0.5 in case of NSFNet topology, the load slightly below 0.5 for Cost239 topology and the load value slightly above than 0.3 in case of 4x4 Torus topology. The losses above these load values are because of the reason that local buffers (BTQs) at highly loaded links reach to their maximum capacity, and these losses are local drops which also increase the burst loss of 1-hop bursts.

Figures 4.23, 4.24 and 4.25 show the burst loss performance based on number of hops for RCIB for all three network topologies respectively. RCIB buffers only those transit bursts which have already traversed equal to or more than half of the network diameter, hence, providing advantage of intermediate buffering to long distant traffic. It can be observed from the plots that in contrast to LHIB, RCIB improves the loss rate of 3-hop and 4-hop traffic. Whereas, the loss rate for 2-hop burst traffic remains same below load value equal to 0.5 in case of NSFNet topology and increase in loss rate of 2-hop traffic in Cost239 and 4x4 Torus network topologies. The rise in the loss rate of 1-hop traffic is due to the overflow from the local buffer in the high loaded scenarios.

The improvement in the fairness among different hop count burst loss can be observed from coefficient of variation for the blocking probabilities shown in Figures 4.26, 4.27 and 4.28. The coefficient of variation (c_v) given in Equation 4.1, is the ratio of standard deviation σ to the mean μ, which gives the unfairness measurement for the entire network in a quantitative way.

$$c_v = \frac{\sigma}{\mu} \tag{4.1}$$

It can be observed that for NSFNet network, no intermediate buffering performs better in terms of fairness below load value 0.7, while above this value, in high loaded situations, LHIB outperforms both RCIB and no intermediate buffering. In case of Cost239 based topology, LHIB performs better in all load scenarios. The reason for this performance is reduction in loss rate of 2-hop burst traffic, as in this topology, most of the traffic is comprised of 2-hops. For the case of Torus network , RCIB remains better than the other two because of reason that there is more traffic for longer paths, and longer path traffic gets the advantage of intermediate buffering in RCIB.

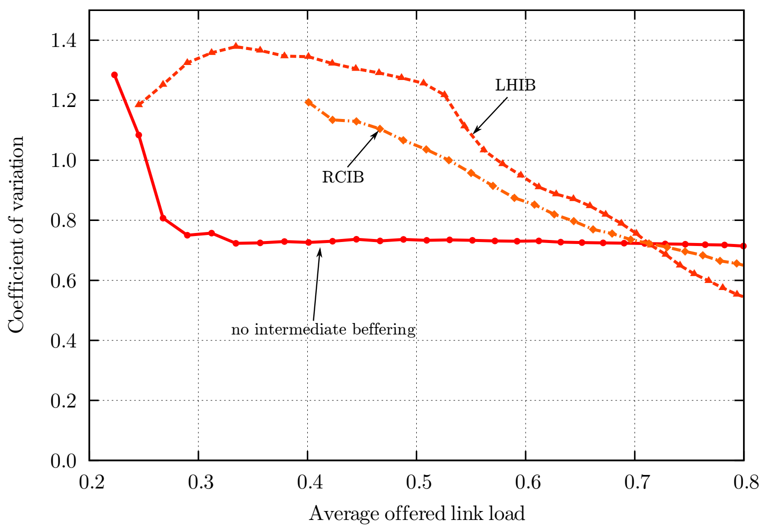


Figure 4.26: CoV for blocking probability versus average offered link load in NSFNet network.

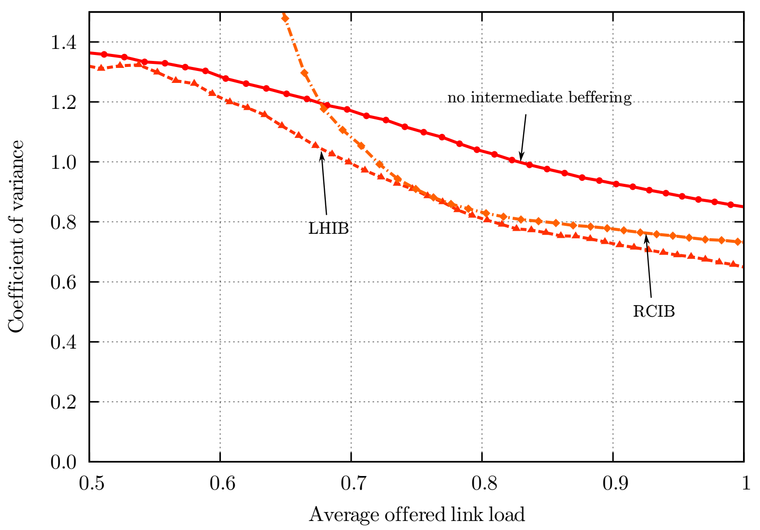


Figure 4.27: CoV for blocking probability versus average offered link load in Cost239 network.

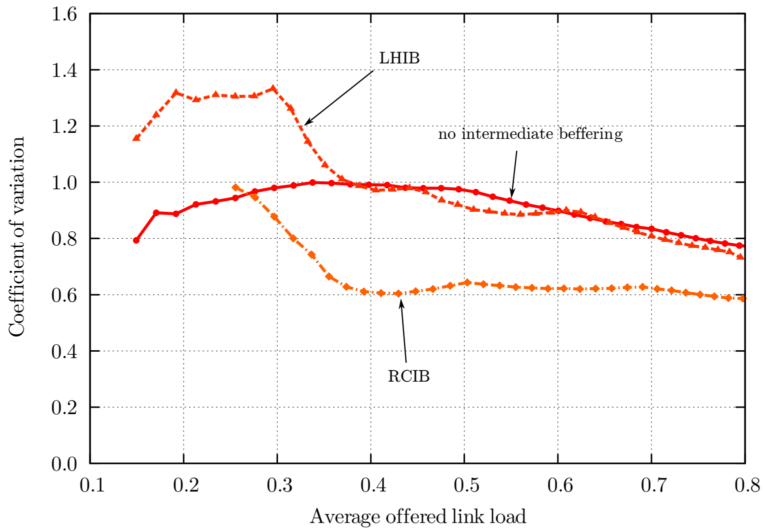


Figure 4.28: CoV for blocking probability versus average offered link load in 4x4 Torus network.

4.5 Summary

In this chapter we have presented the requirement of edge-core joint node (ECJN) in optical burst switched (OBS) networks. Architecture of ECJN is described in detail with functionalities of each module. A new scheduling algorithm, composite edge-core scheduling with void filling (CECS-VF) for multiplexing both local and transit traffic is proposed and the buffer requirements for the ECJN is discussed.

The effect of high transit load on local assembled bursts, and effect of high local load on transit traffic is analyzed through simulations.

Restricted intermediate buffering is proposed, which can be used as a contention resolution approach to improve overall burst loss rate and through simulations it is also shown that intermediate buffering improves in fairness of different hop count burst loss rate.

Chapter 5

Routing Strategies for OBS Networks

OBS networks, in general, use one way resource reservation mechanisms for setting up necessary resources for each burst transmission [Qiao99]. The data bursts are transmitted without the acknowledgement that the resourcs along the path are successfully reserved or not. This mechanism leads to an end-to-end transparent connection. Therefore, whenever the number of simultaneous reservation attempts exceed the number of available resources, some fail. Consequently, due to the lack of sophisticated optical buffers, this results in the burst loss and congestion on central links [Barradas09]. Retransmission of dropped bursts and holding of network resources from ingress node to the dropping node degrades the overall network performance. Therefore, contention in the network represented by the burst loss probability is primary metric of interest in OBS networks.

OBS networks are sensitive to the congestion in most frequently used central links due to this the network becomes overloaded. In overloaded scenarios, the contention resolution mechanisms cannot avoid the collision effectively. Therefore, proactive approaches are preferable to avoid the contention before it occurs. Adaptive routing is one of the proactive approaches which attempts to resolve the contention in space domain before its occurance. Routing strategies exploit the capacity of under utilized links, balance the load on the links, maximize the network utilization and result in reduced burst loss. The added advantage of implementing routing strategies to control congestion is its adaptability to the changes in traffic conditions.

Considerable efforts have been devoted to the study of routing path optimization, path selection, and load balancing in OBS networks. Usually it is very difficult and sometimes impossible to compare the new strategies with the work of other researchers due to the unavailability of complete information of network operating conditions and custom developed simulation environments. As a result, a trend is set in OBS community to compare the results of newly proposed strategies with the shortest path routing. This approach resulted in a useful work, but it became very difficult to analyze and compare those newly proposed routing strategies with each other and conclude to one single solution. In this chapter, we have focused on routing strategies proposed in the literature and some of them are choosen for extensive comparison through simulation using uniform network operating parameters. At the end, a new routing strategy is proposed which remains complete winner in terms of the burst blocking probability and network throughput.

5.1 Classification of Routing Algorithms

In modern communication networks, one of the most important functions of network layer is routing. Routing determines which path in the network is to be taken from source node to the destination node for data transmission with the following objectives [Hendling04] in general and in optical networks in particular:

- Use of shortest feasible path
- Minimize loss
- Balance network load
- Maximize network throughput
- Minimize end-to-end delays

It is hardly possible to achieve all the objectives in one routing strategy. For example, a routing strategy which minimizes the loss, balances network load and maximizes network throughput, achieves these advantages at the cost of a little increase in end-to-end delays. These delays are caused by shifting the load on longer paths to avoid congestion on the shorter paths. Routing strategies use metrics to evaluate which path will be suitable for the bursts to route. Path bandwidth, propagation delay, availability, load, blocking probability, congestion, utilization and link frequency are commonly used metrics for path determination. Routing tables are maintained to store routing information and network state information. It depends on the routing algorithm that what type of routing table and what information is to be stored and what type of information is to be exchanged between network nodes. Routing algorithms can be classified into following categories:

5.1.1 Static or Dynamic

Routing algorithms compute the paths and store them in the routing tables. It depends upon the type of routing algorithm that how to calculate paths and when to modify the routing tables. In static routing algorithms, one or more routes are calculated in advance, based on some static metric such as physical distance or number of hops or a combination of both, and are changed only by administrators. These type of routing algorithms are suitable only for steady traffic networks and cannot be adapted to the changing traffic conditions.

Dynamic or adaptive routing algorithms are adaptive to the current network changes such as traffic load, link congestion, loss rate and topological changes. Dynamic algorithms maintain network state information through periodic status message updates from all nodes. The process of maintaining state information and their usage is categorized as three routing strategies [Hendling04]:

Source Routing

In source routing, all nodes in a network maintain global state information about each link and the network topology. A feasible path is locally determined and each entity uses this information for making decisions. A link state routing protocol is operational to update global state at every node. Source routing is simple to implement, maintain, debug and upgrade and it also guarantees loop free paths. Whereas, the computational cost at each node increases, as periodic status updates aggravate signaling overhead.

Distributed Routing

In distributed routing also called hop-by-hop or isolated routing, decisions are made separately on each node, based only on the local information. Distributed routing algorithms need distance vector protocol to maintain a global state in terms of distance vectors at every node. Routing response time is shorter in distributed routing algorithms, but loops may occur due to incomplete global state information.

Hierarchical Routing

In hierarchical routing, nodes are clustered in groups. These groups are further clustered into high-level groups recursively, creating a multilevel hierarchy [Hendling04]. Every node maintains aggregated global state information with detailed state information about the nodes in the current group and information of other groups. Hierarchical routing decreases the complexity of network topology and signaling overhead and increases routing efficiency.

5.1.2 Single-path or Multi-path

In single-path routing strategy only one path is stored in the routing table for one source destination pair. Single path routing strategies can be static or dynamic. In case of static, only one path is computed between a node pair, and all traffic between this pair is routed on this path. Whereas, dynamic single path strategies compute single paths periodically.

In general, several equally good paths exist between each source destination pair, routing strategies maintain multiple paths between each pair of nodes. Traffic is usually splitted between all paths with a splitting factor or path weighting factor. Multi-path routing strategies better utilize network resources and share the load on all the network links. Also different paths can be used for different classes of traffic. Multi-path routing strategies also increase reliability of networks in case of link or node failures.

Alternative routing or deflection routing is special case of multi-path routing. If one path is unavailable, the traffic is deflected or routed to the alternate path [Klinkowski10].

5.1.3 Link-state or Distance vector

In case of adaptive routing, network state information is to be updated periodically. This update can be based on link state or distance vector

Criteria	Distance vector	Link state
Update	changes communicated to neighbor nodes	communicated to all nodes (i.e., broadcast)
Complexity	count-to-infinity problem and slow convergence, respectively	$\mathcal{O}(\lvert V\rvert \cdot \lvert E\rvert)$ data units required, fast convergence
Shortest path	Bellman-Ford algorithm	Dijkstra algorithm
Robustness	not much robust	robust

Table 5.1: Comparison of distance vector and link state routing protocols.

protocol. Both of these protocols are compared in Table 5.1.

The link state protocol converges fast into a stable state and comparatively reliable, therefore, more suitable for large networks. On the other hand, link state protocols are complex and require more computing resources.

5.2 Route Calculation and Selection in OBS networks

In OBS, routing techniques involve two stages; route calculation and route selection [Thodime03]. Both route calculation and route selection can be static or dynamic. In static route calculation, one or more paths are calculated in advance based on the static network information. The dynamic route calculation techniques recompute paths periodically according to dynamic network behavior in order to adapt to the network dynamic behavior.

If the route calculation phase computes more than one paths either statically or dynamically, the route selection phase selects one of the candidate path based on some defined criteria. Path selection strategy weights the routes according to routing metrics representing the basic and current status of the network links and paths, and are used for routing decisions.

Route selection can either be static or dynamic. In static route selection, traffic is splitted on to multiple paths based on certain static parameters, such as hop count, distance, fixed or variable fraction. Whereas, the dynamic route selection policies are similar to the dynamic route calculation strategies, and the paths are dynamically selected based on the changing network state. On the basis of exchange of link status messages or feedback of each transmission, a cost function is calculated for each route which is used to shift traffic to less loaded routes. Path selection techniques are threshold based, probabilistic and rank based [Klinkowski10]. Threshold based policies shift the traffic on secondary paths if a certain threshold is met. Probabilistic policies split the traffic with certain splitting factor. Rank based, as its name specifies, ranks the paths according to some priority and selects the one with the highest priority.

The status messages updates necessary to make dynamic route calculation and selection can be made in two ways; probe based or broadcast based [Thodime03]. In probe-based, the source node sends probe messages to the core nodes, and core nodes respond with the required status. Feedback of success or failure of each transmission is an example of probe messages, resulting in a higher signaling overhead. In the broadcast based approach, each node broadcasts its status to all nodes periodically based on some interval. In OBS, in order to reduce the signaling traffic, it is better to broadcast status messages in case if there is any significant change in the congestion status of the link.

5.3 Routing in OBS networks

As described in Section 5.2, routing in OBS can be single-path or multi-path, both of which can be static or dynamic. The study of routing in OBS networks is classified into single-path and multi-path routing.

5.3.1 Single-Path Routing

The static single-path routing algorithms are ususally based on shortest hop path, shortest distance path, Dijkstra algorithm and widest shortes path algorithm [Garcia08]. Dynamic single-path routing algorithms recompute the shortest or feasible paths periodically based upon link costs after the network state update. A number of dynamic single-path routing strategies for OBS networks are proposed in the literature. We review some of the routing strategies, and in Table 5.2, these are listed with references, name of the algorithm and the route calculation metric used.

Reference	Algorithm	Calculation metric
[Thodime03]	Congestion based	Link load
[Boudriga03]	WRPA	Link failure probability
[Hyyti04]	OBS-aware	Blocking
[Zhang04a]	-	Link load
[Zhang04b]	-	Link load
[Teng05]	-	Blocking
[Ou05]	MWVT-routing	Blocking
[Huang05b]	DSR	Link utilization
	DAR	Link utilization
[Li06]	BCLB	Link load
[Du06]	ALB-Dijkstra	Link frequency
[Garcia07]	ED	Link frequency
[Chen08]	SLNS-Heur	Blocking
[Barradas10]	SBPR-nPP	Link frequency

Table 5.2: Dynamic single-path routing algorithms in OBS networks.

Thodime proposed in [Thodime03] least-congested dynamic route calculation technique for the route calculation at every connection request. The congestion of each link is set as the weight of the link, and paths are computed based on these weights. It is shown in the results that the propsed algorithm performs better than the shortest path routing.

In [Boudriga03], a wavelength routing predictive algorithm (WRPA) was proposed. This algorithm defines the cost of each link based upon the failure probability of each link. Using these costs, shortest path is determined between each source-destination node pair.

A mixed integer linear programming (MILP) formulation for the routing and wavelength assignment problem in OBS networks was proposed in [Hyyti04]. The resulting configuration is based on the centralized routing and is referred to as OBS-aware routing. The main idea of this routing is that the bursts cannot contend after they have traversed m hops. As a result, fairness is improved among links because bursts have to compete for resources only at first m hops. This also improves the overall efficiency of the network because the later the burst is dropped, the more resources it wastes. The researchers have used MILP-based formulation to solve the problem, which is computationally very complex for large networks and the authors have claimed that better solutions are needed in the future. The results of the proposed solution are compared with the shortest path routing, and improved overall burst loss rate of network.

In [Zhang04a], explicit routing for traffic engineering in labeled optical burst switched WDM networks was proposed. Based on integer linear programming formulation (ILP), this routing approach sets a load threshold referred to as watermark and balances the load on the paths by minimizing the number of links whose load is greater than the watermark. Numerical results showed that the three proposed optimization approaches can significantly reduce burst loss probability without too much sacrifice in the average hop distance for a burst.

In [Zhang04b], a pre-planned global rerouting for fault management in labeled optical burst switched WDM networks was proposed. Three integer linear programming (ILP) formulations are used in which both the normal routing table, used under normal operation, and the backup routing tables, used after failure occurs, are pre-optimized. The objective is to achieve an optimal load balancing both before and after a failure such that the network state can still remain stable with the minimum burst loss probability when the failure occurs. The proposed formulations are compared with the shortest path routing and the improvement in the burst loss probability is proved.

A centralized routing and traffic engineering based routing path optimization in optical burst switched networks was proposed in [Teng05] with the objective of balancing the traffic across the network links in order to reduce the congestion and improve the overall performance. The researchers used linear programming (LP) formulations to determine all possible shortest paths between each source-destination pair, and an heuristic technique is applied to assign a single path to route the bursts between each source-destination pair. The second approach is an integer linear programming (ILP) based solution to the same formulation to determine a single path, so that the overall network loss is minimized. The authors mentioned that the ILP-based solution is applicable to only moderate size networks, for bigger networks with hundreds of nodes and links, it is not possible to obtain an optimal solution within a reasonable amount of time. Both of the proposed approaches are compared with the shortest path routing and yielded improvement in overall burst loss rate.

A study on dynamic load balanced routing techniques in time-slotted optical burst switched networks was presented in [Ou05]. In this research, the shortest-cost path calculation is implemented using link weight functions that depend upon current specific network load information on each wavelength. After each routing table update, the shortest path is computed and stored in the routing tables at each node. The link weight function is based on the congestion and hop distance to the destination. In simulation, three policies were considered and compared which are hop-only routing, capacity-only routing and hop and capacity routing.

Two dynamic routing algorithms, dynamic routing with synchronous rerouting (DSR) and dynamic routing with asynchronous rerouting (DAR) were proposed in [Huang05b]. In both of the proposed algorithms, link weights are calculated based on the link utilization and length of the link. Using these link weights as costs of the links, shortest paths are computed periodically after each status update in DSR algorithm. Whereas, in DAR, the path in the routing table is deleted after a specific time, and recomputed at the connection request. Results prove that the proposed algorithms perform better in terms of the burst loss probability as compared to the static shortest path algorithm.

In [Li06], burst cloning with load balancing (BCLB) scheme was proposed which combines the basic cloning technique with the load balanced routing for cloned bursts. Two paths are calculated between each source-destination pair, first one is active path (AP) for original bursts and second one is backup path (BP) for cloned bursts. The original bursts are routed on APs, whereas the cloned bursts are routed over BPs. APs are usually shortest paths and to avoid hot spots, load balancing is also considered for APs. If original bursts contend with the cloned bursts in the core network, the priority is given to original bursts and cloned bursts are preempted. Integer linear programming (ILP) is used to formulate the proposed routing scheme. Numerical results are compared with the shortest path routing and basic burst cloning (BBC) proposed in [Huang05a], and an improvement in the burst loss probability is shown.

An adaptive load balancing routing algorithm, called ALB-Dijkstra was proposed in [Du06]. This algorithm assigns initial cost of each link equal to 1, and finds shortest paths using ordinary Dijkstra algorithm with increasing cost of each link when used in a path. At the completion of all the paths between every node pair using the proposed method results in a complete set of optimized routing paths. The proposed algorithm is compared with the shortest path routing and an improvement in terms of the burst drop probability is shown with different wavelenths.

An extension to Dijkstra algorithm was proposed in [Garcia07]. The extended dijkstra (ED) algorithm first computes the shortest paths using Dijkstra algorithm, and the ED is applied in case if there exist possible equal cost routes between any pair of nodes. The proposed algorithm computes fully balanced and symmetric routes between all source-destination pairs. Results of ED are compared with original Dijkstra, and an improvement in the burst loss probability is shown.

In [Chen08], route optimization in optical burst switched networks considering the streamline effect was proposed. Researchers in this article proposed two route optimization problems. The first problem, normal state route (NSR) optimization 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: failure recovery route (FRR) optimization, considers the failures, and then primary and backup paths are determined in such a way to minimize the expected burst loss over the normal and failure states. These formulations were based on mixed integer linear programming (MILP) formulations, since the MILP-based formulations are computationally intensive, heuristic algorithms were developed for the problem solution. The two heuristics are streamline effect based normal state route optimization heuristic (SLNS-Heur) and streamline effect based failure recovery route optimization heuristic (SLFR-Heur). The results of proposed heuristic algorithms are compared with the widest shortest path (WSP) and the shortest path first (SPF) algorithms. The results presented in [Chen06] are similar to that in [Chen08].

Streamline based pre-planned routing with no pre-determined paths (SBPR-nPP) was proposed in [Barradas10]. An integer linear programming (ILP) formulation is given for the route calculation. The path calculation for each source-destination is based on two factors, number of hops and link frequency factor that is defined as the number of routes competing for an output link. The proposed strategies perform well on networks with full wavelength conversion capability as compared to the shortest path routing. Whereas, in case of no wavelength conversion availability, there is not any significant improvement as compared to the shortest path routing.

5.3.2 Multi-Path Routing

Dynamic multi-path routing strategies involve path calculation and selection. Usually, two or more paths are calculated statically in advance, and on every connection request, one of the already calculated path is selected according to a defined criteria. Some of the dynamic or adaptive multi-path routing strategies from the literature are reviewed and listed in the Table 5.3 with references, name of the algorithms, path selection criteria and path selection metric used in the given strategy.

Thodime et al. in [Thodime03] proposed congestion based static route calculation technique for load balanced routing in optical burst switched networks. This congestion based routing technique statically computes link-disjoint alternate paths between each source-destination pair, and dynamically selects one of the paths based on the collected congestion information. This technique uses maximum load threshold, which, if exceeded, will signal congestion on the link. Using this congestion value, it calculates congestion status of both primary and alternate paths, and routes the bursts on the less congested path. Simulation results prove that the proposed strategy outperforms the shortest path fixed routing at the cost of slight increase in end-to-end delay and extra signaling traffic.

In [Ganguly04], a multi-path adaptive optical burst forwarding method was proposed for load balancing in optical burst switched networks. The decision to select a path among multiple paths for the same destination is taken locally based on the local knowledge of burst drop rates on the paths. The researches also proposed a multi-path burst pipelined forwarding to resolve the out-of-order packet delivery caused by the multi-path routing. It is proved in the results that the adaptive routing improves burst loss rate as compared to non-adaptive routing at the cost of increase in signaling

Reference	Algorithm	Route selection	Selection metric
[Thodime03]	Congestion based	Rank-based	Link load
[Li04]	AARA	Probabilistic	Blocking
[Ganguly04]	-	Rank-based	Blocking
[Ishii05]	Self-learning	Rank-based	Path priority
[Yang06]	WBLU	Rank-based	Utililization
	WLC	Rank-based	Blocking
	EPP	Rank-based	Path priority
[Hirota06]	-	Rank-based	Path priority
[Lu06]	-	Probabilistic	Link load
[Argos07]	AMOR	Probabilistic	Blocking
[Pedro07b]	-	Probabilistic	Link load
[Klinkowski07a]	-	Probabilistic	Blocking
[Gao08]	Hop-FCR	Rank-based	Path priority
	Hop-LC	Rank-based	Path priority
	Hop-N-FCR	Rank-based	Path priority
[Levesque09]	GPRM	Rank-based	Blocking
[Gao09]	Hop-LC-CC	Rank-based	Path priority
[Diaz09]	RPBS	Rank-based	Path priority
[Barradas09]	MCL	Rank-based	Link frequency
	MEC	Rank-based	Link frequency
[Barradas10]	SBPR-PP	Rank-based	Link frequency
[Triay10]	ACRWA	Rank-based	Path priority

Table 5.3: Dynamic multi-path routing algorithms in OBS networks.

traffic.

A new contentionless dynamic routing protocol for OBS using wavelength occupation knowledge was proposed in [Agusti04] to combat the contention problem in OBS. Based on the slotted wavelength occupation knowledge (SWOK), in this approach, wavelength assignment, routing and time-slot assignment subproblems are treated jointly. At every node a wavelength occupation vector is maintained for all outgoing links and a dynamic offset window mechanism is defined to prevent conflicts of segments of path-sharing by more than one sources. Combining these functions with path selection and wavelength assignment, dynamic routing protocol is achieved.

A dynamic load balancing in ip-over-wdm optical burst switching networks was proposed in [Li04]. Two algorithms, shortest-hop path routing (SHPR) and widest-shortest-hop path routing (WSHPR) were proposed for route calculation, and adaptive alternate routing algorithm (AARA) was proposed for path selection based on the probability of splitting traffic to already calculated link-disjoint paths. The proposed algorithm works in a dynamic way, so that in each time window, the traffic is distributed using a traffic shift probability which is calculated by the feedback messages from all the nodes at the end of every time window. Through simulations, the authors have proved that their proposed strategy performs better than the shortest-hop path routing-fixed routing (SHPR-FR) and static alternate routing (SAR) in terms of the burst loss probability. Also the authors have verified their algorithm using different traffic scenarios.

A self-learning route selection scheme using multi-path searching packets in an OBS network was proposed in [Ishii05]. Initially, the routing algorithm sets equal priority of each path and with every burst transmission there is a feedback message of either success or failure. Priorities of all paths are updated using feedback messages, and an edge router sends a burst on the route that has the highest priority of among all the routes. Results prove that the proposed algorithm performs better than the conventional shortest path routing algorithm.

Lu et al. in [Lu06] proposed a gradient projection based multi-path routing algorithm (GPMR). Using linear programming formulation, this algorithm splits the traffic among multiple paths between each source destination pair. The key idea of this algorithm is to let the source node of each source destination pair periodically measure the offered load on the links that are traversed by its alternative paths. Then at the end of each time window, the source node calculates each path's first derivative length to evaluate the variation of the impact of the offered burst traffic on the path. Based on the above information, the gradient projection algorithm will be applied to obtain the amount of burst traffic that will be offered to each alternative path for the next time window. Through simulations, it is shown that the proposed algorithms improve the burst loss probability in comparison with already proposed adaptive alternate routing algorithm (AARA) proposed in [Li04].

In [Hirota06], a cooperation method of routing and wavelength assignment is proposed for adaptive routing and load balancing in OBS networks. In this algorithm, routes are calculated statically, and at each node a suitability index is maintained for every wavelength. At the time of path selection, this suitability index is used to rank the best path. With every transmission, either successful or not, a feedback message is forwarded to the source node to update the suitability index. Routing decisions are made locally at each node considering the local information only. The bursts destined to the adjacent nodes are routed differently using the least-used algorithm, hence saving the bandwidth for the future requests. Through simulations, the researchers have shown that their proposed algorithm performs better than the shortest path routing with random wavelength assignment and shortest path routing with priority-based wavelength assignment.

In [Yang06], the authors have proposed three routing strategies for adaptive path selection, weighted bottleneck link utilization (WBLU), weighted link congestion (WLC) and end-to-end path priority-based (EPP). All of these three strategies calculate path statically based on number of hops or link distances, and at the time of connection request, select one of the pre-calculated paths dynamically. The first one, WBLU selects less utilized paths, and therefore, performs well in low loaded situations. WLC strategy calculates the path blocking probability using congestion status of the links traversed by the path and selects the path with less probability of contention. Both of these strategies also take into account the number of hops while making path selection. The third strategy EPP priotizes the paths through feedback messages of bursts success or failure in the transmission. Each successful transmission increases the priority of the path and vice versa, dropped burst decreases the priority. The strategy selects the path with the higher priority while transmitting the bursts. All of these three proposed strategies behave differently in different traffic conditions and topologies. Therefore, the authors suggested majority binary voting (MBV), weighted nonbinary voting (WNV) and dynamic weighted nonbinary voting (DWNV) strategies to select one of the path out of three paths selected by all three path selection strategies. As a result, DWNV remains complete winner in all load scenarios. At every connection request for the burst arrival, this policy has to make path selection for three routing path selection strategies and then convergence to the optimal one would increase complexity of the path selection strategy. Moreover, the third path selection strategy EPP is based on the feedback messages. These feedback messages for every burst can lead to increased signaling overhead in the network.

A non-linear optimization for multi-path source routing in OBS networks was proposed in [Klinkowski07a]. In this routing scheme, multiple paths are computed between each source destination pair. Using a gradient optimization method, they calculate a traffic splitting vector that determines the distribution of traffic over these paths by taking into account the burst blocking probabilities on the links. Through simulations, it is proved that the proposed algorithm performs better than the shortest path routing and pure deflection routing (DR). The researchers also claimed that the proposed non-linear optimization method is the avialable alternative for the linear programming formulations based on piecewise linear approximations.

In [Pedro07b], efficient multi-path routing for optical burst switched networks was proposed to the reduce congestion on the bottleneck links. Two linear programming models were proposed, first only minimizes the burst traffic load on the most congested link, whereas the second model fine tunes the solution of the first model such that the average offered traffic load per link is also reduced. The proposed method is based on a centeralized routing which computes the candidate paths in advance, and through linear programming formulation, fraction of the burst traffic is routed through each candidate path. Moreover, the researchers also proposed the transmission delay for ordered bursts delivery through different paths. Results show that the multi-path routing improves the burst loss probability as compared to the single-path routing.

The researchers in [Argos07] proposed the adaptive multi-path OBS routing (AMOR) algorithm to obtain the better performance in OBS networks regarding burst loss probability and load balancing. This strategy initially routes the burst using equal cost multi-path (ECMP) with number of hops metric, and later, after periodic congestion information, calculates link cost functions and performs load distribution based on these link costs and the link loads status. While calculating the link cost function, it also considers the number of hops to the destination for the path comprising of this link. The drawback of this strategy is that, instead of considering the blocking probability of highly congested link in the path, it only takes into account the current output link, although the authors have claimed that the link state information is available locally. The decision based on local output links of source node may not be enough and the strategy can select the paths comprising of highly congested centeral links of the network.

The authors in [Gao08] presented three fairness improving adaptive routing algorithms for the contention resolution and improving the burst loss probability for JET based OBS networks using the latest available unscheduled channel (LAUC) for wavelength scheduling. The three proposed algorithms are based on hop-by-hop routing using forward channel reservation (Hop-FCR), hop-by-hop routing using link connectivity (Hop-LC) and hop-by-hop routing using neighborhood forward channel reservation (Hop-N-FCR). The researchers improved the burst loss probability and resolved burst loss unfairness between longer paths and shorter paths in terms of number of hops. Distributed routing when implemented in OBS networks, complicates offset time setting issues, as offset times of the bursts cannot be changed at the intermediate nodes without using fiber delay lines if bursts paths are to be changed to the longer paths due to the link preferences. The authors have not addressed offset time issues created with distributed routing. The simulations are performed on 5x5 mesh-torus topology, which is a balance topology, the algorithms performance on networks with less nodal-degree is unknown.

To reduce the negative impact of wavelength conversion cascading constraints, the authors in [Gao09] proposed a proactive routing scheme. The routing scheme uses hop-by-hop routing using forward channel reservation method considering the routing and wavelength reservation problems jointly. Unlike their previous work [Gao08], latest available unscheduled channel with void filling (LAUC-VF) is used for channel scheduling, and results are compared with fixed shortest path routing, fixed shortest path under conversion cascading constraint, and fixed shortest path routing with conversion avoidance. It is shown that their proposed algorithms improve burst loss probability and fairness among different hop count routes.

Graphical probabilistic routing model (GPRM) was proposed in [Levesque09]. In this research, the authors proposed a routing scheme that selects less utilized links by using a bayesian network. The routing table is periodically updated by using the bayesian network and decisions are made on local knowledge and there exists one decision hop per possible next hop. This routing scheme uses distributed routing, where decision of next hop is taken at each node using local output links state. Offset time, burst loss rate, number of hops to the destination and destination identifier are used for the decision. Through simulations using network simulator 2 (ns-2) with 14 node NSFNet based topology, it is shown that proposed algorithm (GPRM) performs better than the shortest path algorithm in terms of burst loss rate and throughput with negligible increase in end-to-end delay.

Routing with prioritization based on statistics in OBS networks (RPBS) was proposed in [Diaz09]. With acknowledgement of each burst transmitted, a path priority is maintained, and at the time of path selection the path with higher priority is selected to route the burst. The researchers also claim that the proposed strategy is QoS aware by making size dependent burst assembly and offset times. Moreover, a waiting queue is introduced at every edge router, where a copy of each transmitted burst is queued. Upon acknowledgement, retransmission is made from waiting queue if burst is blocked in the core part of the network.

Edge-node deployed routing strategies for load balancing in optical burst switched networks were proposed in [Barradas09]. Linear programming optimization technique was used for pre-calculation of eligible paths and two strategies, minimizing the maximum congested link (MCL) and minimizing the maximum end-to-end congestion (MEC) were proposed for path selection. MCL is based on the idea that more a certain link is included in the chosen paths for source-destination pairs, the higher its blocking probability can be. Therefore, paths for the source-destination pairs should be selected with the objective of minimizing the blocking probability of the link. Whereas, MEC is based on the idea that the blocking may occur at any link traversed by a burst along the path. Therefore, paths for the source-destination pair should be selected so that demands have the lowest probability of contending with other demands at every link from the source to destination, minimizing end-to-end blocking. Simulations prove that the proposed strategies perform better than the shortest path routing.

Streamline based pre-planned routing using K pre-determined paths (SBPR-PP) was proposed in [Barradas10]. An integer linear programming (ILP) formulation is given for the selection of path from already calculated set of paths. The path selection is based on two factors, number of hops and link frequency factor that is defined as the number of routes competing for an output link. The proposed strategies perform well on networks with full wavelength conversion capability as compared to the shortest path routing. Whereas, in case of no wavelength conversion availability, there is not any significant improvement as compared to the shortest path routing.

Triay and Pastor [Triay10] proposed an ant-based distributed routing and wavelength assignment algorithm for OBS networks. The algorithm is based on ant colony optimization to adapt better to the network changes in real time and computes pheromone concentrations according to the congestion level of the link and path length in terms of number of hops. The pheromone table is later updated by negative feedback in case of blocked burst and positive feedback in a result of successful transmission. The decision of next hop is taken at evey node based on the local information. It is shown that the proposed algorithm outperforms shortest path routing with random wavelength selection, random path routing and wavelength assignment and shortest path routing with first-fit traffic engineering wavelength selection protocols at the expense of heavy control traffic.

5.4 Selected Routing Strategies

A properly designed routing strategy is able to enhance the network performance by reducing the congestion and the burst loss rate, balancing load on links, increasing network throughput and better utlization of network resources. In OBS, the major problem is contention in the core network, therefore, routing strategies are used as one of the methods to reduce contention.

Static single-path routing algorithms are more prone to burst contention because they are not able to change the route in case of congestion. Whereas, dynamic single-path routing strategies are able to recompute the paths if there exists congestion or path blocking probability is higher. Periodic calculation of paths yields better results, but this methodology may be computationally expensive until and unless the centralized routing is not implemented, which is difficult to implement in large scale networks.

Dynamic multi-path routing with static route calculation and dynamic selection is computationally less complex because the path calculation is to be performed in the offline mode, and at the time of each burst arrival, selection of feasible path out of pre-calculated paths is an efficient procedure. Therefore, dynamic multi-path routing algorithms are selected for study. The path calculation and path selection can be performed either using mathematical based LP, ILP and MILP to solve optimization functions in search of best paths that minimize congestion, or heuristic algorithms can be used to route the bursts in a way to reduce the congestion in the network. Also, heuristic algorithms are less complex and scalable while MILP algorithms are complex and not scalable [Coulibaly11]. Therefore, heuristic methods for the path calculation and selection are considered in this work.

From Table 5.3, we can realize that all of the dynamic path selection strategies proposed in the literature use burst blocking probability, link load, link utilization, path priority and link frequency to rank the paths. All of these metrics are used differently to weight the links, calculate the cost function of the links and path selection. We have chosen the following strategies:

- Congestion-based technique – based on link loads
- WBLU – based on link utilizations
- WLC – based on path congestion

The routing strategies maintaining path priorities are not considered because of the fact that they require acknowledgement messages after each burst transmission causing higher signaling overhead. Moreover, the strategies using link frequencies as the weighting factor were based on optimization techniques. At the end, a new routing strategy, weighted link congestion and frequency (WLCF) is proposed embedding link frequency with path congestion to find the optimal path.

5.4.1 Congestion-Based Technique

Notations used in this technique are given in Table 5.4

Symbol	Description
$e_{(i,j)}$	link between node i and node j in the network
ρ_e	load on link $e_{(i,j)}$
ρ_{max}	maximum load threshold, which if exceeded, will signal congestion on the link
$LS_{(i,j)}$	load status of the link $e_{(i,j)}$.
$LS^p_{(s,d)}$	load status of primary path
$LS^a_{(s,d)}$	load status of alternate path
$r^p_{(s,d)}$	primary path between source-destination pair (s,d)
$r^a_{(s,d)}$	alternate path between source-destination pair (s,d)

Table 5.4: List of notations used in congestion based technique.

Congestion based dynamic route selection technique using fixed alternate shortest paths, porposed in [Thodime03], statically computes link-disjoint alternate paths and dynamically selects one of the paths based on the collected congestion information. When maximum link threshold ρ_{max} is exceeded, it signals the congestion on the link. Therefore, if the load on link $e_{(i,j)}$ is above the threshold value, that is, $\rho_{(i,j)} \geq \rho_{MAX}$, then the load status of the link is set to one. Once the load status of all the links of a node is determined, this information is sent to all nodes in the form of load status packet.

When the edge node has a burst ready to be transmitted, the node determines whether the burst has to be sent on the primary path or on the alternate path. The edge node calculates the load status of the primary path using Equation 5.1.

$$LS^p_{(s,d)} = \sum_{e_{(i,j)} \in r^p_{(s,d)}} LS_{(i,j)} \tag{5.1}$$

where, $r^p_{(s,d)}$ represents primary path between source s and destination d.

If the load status of the primary path is greater than zero, then at least one of the links in primary path has its load status set to one, indicating the congestion on primary path. In this case, the edge node calculates load status of alternate path using Equation 5.2.

$$LS^a_{(s,d)} = \sum_{e_{(i,j)} \in r^a_{(s,d)}} LS_{(i,j)} \tag{5.2}$$

If load status of alternate path is zero, then the burst is transmitted on the alternate path, otherwise, the burst is sent on the least congested path, i.e., the path corresponding to $min(LS^p_{(s,d)}, LS^a_{(s,d)})$. A high level view of congestion based technique for path selection is given in Figure 5.1.

Congestion Based Route Selection phase:
INPUT: Path matrix and the request $r(s, d)$.
OUTPUT: The index of path from s to d with minimum congestion.

1. Calculate link status for all links.
2. Calculate load status of primary path using Equation 5.1.
3. If $LS^p_{(s,d)} = 0$ then select primary path.
4. If $LS^p_{(s,d)} > 0$ then calculate load status of alternate path using Equation 5.2.
5. If $LS^a_{(s,d)} = 0$ then select alternate path.
6. If $LS^a_{(s,d)} > 0$ then select path corresponding to $min(LS^p_{(s,d)}, LS^a_{(s,d)})$.

Figure 5.1: High level view of Congestion Based Technique.

5.4.2 Weighted Bottleneck Link Utililization (WBLU) Strategy

The WBLU strategy [Yang06] maintains the link utilization information and uses this information to rank the cadidate paths. The motivation behind this strategy is to reduce or prevent the contention by using paths with less utilized links. Table 5.5 contains the list of notations used in WBLU routing strategy.

Symbol	Description
s	source node $s \in V$
d	destination node $d \in V$
$e_{(i,j)}$	link between node i and node j in the network
W_e	number of wavelengths in link e
$U(e,t)$	utilization of link e at time t
N_s	number of successful bursts
T_i^s	length in time of ith successful burst
k^*	selected index of path π
π_k	kth path between a source-destination pair
K	number of candidate paths from node s to d
n_k	number of hops in kth path

Table 5.5: List of notations used in WBLU strategy.

The utilization of a link e at time t is defined as:

$$U(e,t) = \frac{\sum_{i=1}^{N_s} T_i^s}{W_e \times t} \tag{5.3}$$

Where W_e is the number of available wavelengths; at time $t = 0$, we assume that the utilization $U(e, 0) = 0$ for all links e. T_i^s is length in time of ith successful burst and N_s is total number of successful bursts.

The weighted bottleneck link utilization (WBLU) strategy routes bursts from s to d along the path $\pi_{k^*(t)}$ whose index $k^*(t)$ is obtained using the following metric:

$$k^*(t) = max_{1 \leq k \leq K} \frac{1 - max_{1 \leq i \leq n_k} U(e_i, t)}{n_k} \tag{5.4}$$

The numerator in Equation 5.4 is the available capacity of the bottleneck link in a given path k. Therefore, the WBLU strategy routes bursts along the path with the highest ratio of available bottleneck link capacity to path length. By taking the number of hops into account, it is ensured that if the bottleneck link utilization is similar for two paths, then the shortest path is selected for routing; the longer path is selected only if the utilization of its bottleneck link is significantly lower than that of the shorter one. The high level view of WBLU algorithm is given in Figure 5.2.

WBLU Route Selection phase:
INPUT: Path matrix and the request $r(s, d)$.
OUTPUT: The index of least utilized path from s to d.

1. Calculate link utilizations using Equation 5.3.
2. Find the index k^* of the least utilized path from K available paths using Equation 5.4.
3. Route the burst on selected path.

Figure 5.2: High level view of WBLU.

5.4.3 Weighted Link Congestion (WLC) Strategy

The objective of the WLC strategy [Yang06] is to route bursts along the path that is most likely to lead to a successful transmission. The source uses information on link congestion along each path to infer the burst drop rate of the path. A high level view of WLC is given in Figure 5.3. The notations used in this strategy are given in Table 5.6.

Symbol	Description
s	source node $s \in V$
d	destination node $d \in V$
$e_{(i,j)}$	link between node i and node j in the network
$c_{(e,t)}$	congestion level of link e at time t
$N_{succ(e,t)}$	number of successful bursts on link e at time t
$N_{drop(e,t)}$	number of dropped bursts on link e at time t
$Pb_{\pi_k,t)}$	probability of blocking on kth path at time t
k^*	selected index of path π
π_k	kth path between a source-destination pair
K	number of candidate paths from node s to d
n_k	number of hops in kth path

Table 5.6: List of notations used in WLC strategy.

The congestion level of a link e at time t is given as:

$$c_{(e,t)} = \frac{N_{drop(e,t)}}{N_{drop(e,t)} + N_{succ(e,t)}} \tag{5.5}$$

We assume that at time $t = 0$, the congestion $c_{(e,0)} = 0$ for all links e.

Assuming that link drop probabilities are independent, at time t the probability that a burst will be dropped along this path can be calculated as:

$$Pb_{(\pi_k,t)} = 1 - \prod_{1 \leq i \leq n_k} (1 - c_{(e_i,t)}) \tag{5.6}$$

The weighted link congestion (WLC) strategy routes bursts from s to d along the path $\pi_{k*(t)}$ whose index $k^*(t)$ is obtained using the following metric:

$$k^*(t) = max_{1 \leq k \leq K} \frac{1 - Pb(\pi_k, t)}{n_k} \tag{5.7}$$

WLC Route Selection phase:
INPUT: Path matrix and the request $r(s, d)$.
OUTPUT: The index of path from s to d with minimum congestion.

1. Calculate link congestions using equation 5.5.
2. Calculate burst drop probabilities for all paths between $s - d$ pair using equation 5.6.
3. Find the index k^* of the least congested path from K available paths using Equation 5.7.
4. Route the burst on selected path.

Figure 5.3: High level view of WLC.

5.4.4 Weighted Link Congestion and Frequency (WLCF)

Finally, we present a new and simple routing path selection approach, which is based on link congestion and frequency of usage in primary paths. The weighted link congestion and frequency (WLCF) strategy routes the bursts along the paths having less probability of contention by taking into account the path congestion, number of hops and critical links. A link e is more critical if more ingress-egress pairs are using this link in their primary shortest routes. The number of routes that use any given link e in the network is defined as the frequency F_e of the link e. The link with the high

Symbol	Description
s	source node $s \in V$
d	destination node $d \in V$
$e_{(i,j)}$	link between node i and node j in the network
$c_{(e,t)}$	congestion level of link e at time t
$N_{succ(e,t)}$	number of successful bursts on link e at time t
$N_{drop(e,t)}$	number of dropped bursts on link e at time t
$Pb_{\pi_k,t)}$	probability of blocking on kth path at time t
k^*	selected index of path π
π_k	kth path between a source-destination pair
K	number of candidate paths from node s to d
n_k	number of hops in kth path
F_e	frequency of uasge of link e as the number of routes which use link e
α	a weighting factor to assign relative weight to number of hops and frequency

Table 5.7: List of notations used in WLCF strategy.

frequency has higher carried traffic, and gets congested earlier. Moreover, taking into account the frequency of the link keeps as many links as possible available for future reqeuests, as a result indends to avoid congestion. A weighting factor α is used to assign a relative weight to the number of hops and the maximum frequency of the links from the source to the destination while selecting the best among the available paths. The notations used in this strategy are given in Table 5.7.

The congestion level of a link e at time t is given as:

$$c_{(e,t)} = \frac{N_{drop(e,t)}}{N_{drop(e,t)} + N_{succ(e,t)}} \tag{5.8}$$

We assume that at time $t = 0$, the congestion $c_{(e,0)} = 0$ for all links e.

Assuming that the link drop probabilities are independent, at time t the probability that a burst will be dropped along this path can be calculated as:

$$Pb_{(\pi_k,t)} = 1 - \prod_{1 \le i \le n_k} (1 - c_{(e_i,t)}) \tag{5.9}$$

The weighted link congestion and frequency (WLCF) strategy routes the bursts from s to d along the path $\pi_{k*(t)}$ whose index $k^*(t)$ is obtained using the following metric:

$$k^*(t) = max_{1 \leq k \leq K} \left[\alpha \times \frac{1 - Pb(\pi_k, t)}{n_k} \right] \left[(1-\alpha) \times \frac{1 - Pb(\pi_k, t)}{max_{1 \leq e \leq n_k} F_e} \right] \quad (5.10)$$

The weighting factor α gives a relative weight to the hop count and the frequency of the highest value link along the path to the destination. Bigger value of α gives more weight to number of hops and selects the path with a less number of hops to the destination if the congestion of two paths is equal. Whereas, a smaller value of α gives more weight to the frequency of links and selects the path with less frequently used links. With the study of WLCF strategy, we have following observations about weighting factor α.

- Higher value of α gives better results in more meshed networks.
- Smaller value of α gives better results in less meshed networks.

A high level view of WLCF algorithm is given in Figure 5.4.

WLCF Route Selection phase:
INPUT: Path matrix and the request $r(s, d)$.
OUTPUT: The index of path from s to d with minimum congestion.

1. Calculate link congestions using equation 5.8.
2. Calculate burst drop probabilities for all paths between $s - d$ pair using equation 5.9.
3. Find the index k^* of the least congested path from K available paths using Equation 5.10.
4. Route the burst on selected path.

Figure 5.4: High level view of WLCF.

5.5 OBS Network Modelling

We model the network as a graph $G = (V, E)$, where V denotes the set of nodes (i.e., routers) and E denotes the set of edges (i.e., links). Link $e \in E$ comprises W_e wavelengths. The notations used in remaining chapter are described in Table 5.8.

Symbol	Description
s	source node $s \in V$
d	destination node $d \in V$
$e_{(i,j)}$	link between node i and node j in the network
W_e	number of wavelengths in link e
π_k	kth path between a source-destination pair
K	number of candidate paths from node s to d
k^*	selected index of path π
n_k	number of hops in kth path
$c_{(e,t)}$	congestion of a link e at time t
$N_{succ(e,t)}$	number of bursts successfully scheduled along link e upto time t
$N_{drop(e,t)}$	number of bursts dropped along link e upto time t
$Pb_{(\pi_k,t)}$	burst drop probability along path π_k at time t
T_i	length in time of ith burst
F_e	frequency of uasge of link e as the number of routes which use link e

Table 5.8: List of notations used.

5.5.1 Calculation of Link Loads

In an OBS network, the traffic offered to a link e is approximated as the sum of traffic offered to all the paths using this link reduced by the traffic blocked on previous links in the path. The authors in [Rosberg03] called it reduced load approximation that takes into account only those bursts which arrived on the link. In the following work, the offered load ρ_e to a link e is defined as follows:

$$\rho_e = \frac{\sum_{i=1}^{N} T_i}{W_e \times t} \qquad e \in E \tag{5.11}$$

where T_i is the length of ith burst arrived on link e, N is the total number of bursts arrived until time t and W_e is the number of wavelengths on the link e.

Average offered load to a link is given as:

$$\rho_{mean} = \frac{\sum_{e=1}^{E} \rho_e}{E} \tag{5.12}$$

where E is total number of links in the network.

5.5.2 Probability of Blocking

The simplified measurement of the burst loss rate can be the ratio of number of dropped bursts to the number of total bursts arrived on a certain link. Since the bursts of large size have a tendency of facing more contentions which results in dropping of larger chunk of packets. Therefore, the burst blocking probability Pb_e on link e is given as the ratio of time the system remained blocked (time length of blocked bursts) to the total time (time length of arrived bursts on link e).

$$Pb_e = \frac{\sum_{i=1}^{N_d} T_i^d}{\sum_{i=1}^{N} T_i} \qquad e \in E \tag{5.13}$$

where, T_i^d is length in time of the ith dropped burst, N_d is the number of dropped bursts, N is the total number of arrived bursts on the link e.

The mean network blocking probability Pb_{mean} is given as:

$$Pb_{mean} = \frac{\sum_{e=1}^{E} (Pb_e \times \rho_e)}{\sum_{e=1}^{E} \rho_e} \tag{5.14}$$

where E is total number of links in the network.

5.5.3 Network Throughput

In communication networks, throughput or network throughput is the average rate of successful message delivery over a communication channel. In OBS networks, throughput is the average rate of successfully scheduled bursts. The throughput γ_e of link e is given as:

$$\gamma_e = \frac{\sum_{i=1}^{N_s} T_i^s}{t} \tag{5.15}$$

where, N_s is the total number of successfully scheduled bursts in time t and T_i^s is the length in time of the ith successfully scheduled burst.

Mean network throughput is given as:

$$\gamma_{mean} = \frac{\sum_{e=1}^{E} \gamma_e}{E} \tag{5.16}$$

where E is total number of links in the network.

5.6 Simulations and Results

The performance of selected routing strategies described in Section 5.4 is eveluated and compared using the simulation environment discussed in Chapter 3. The simulations were performed on NSFNet based 15 node 25 links topology shown in Figure 5.5, Cost239 based 11 node 26 links topology shown in Figure 5.6, and 16 node 32 links 4x4 Torus topology shown in Figure 5.7. All of the links in given topologies are bidirectional links.

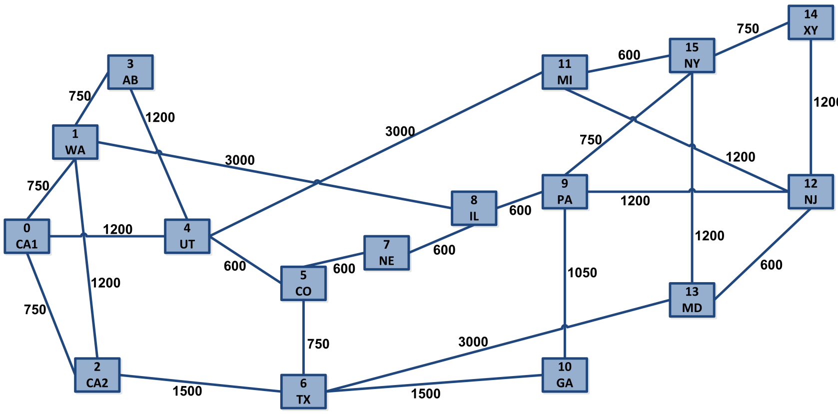


Figure 5.5: NSFNet based 15 node 25 bidirectional links topology

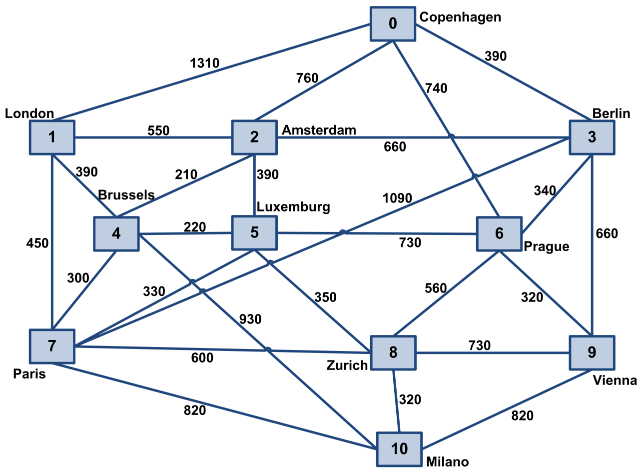


Figure 5.6: Cost239 based 11 node 26 bidirectional links topology

The numbers on the links show the physical distance between the nodes in

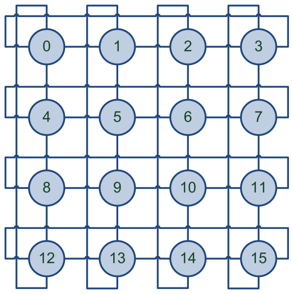


Figure 5.7: 16 node 32 bidirectional links 4x4Torus topology, link length=10 km

kilometers. These distances are chosen for simulation purposes, and do not represent the actual distances of the network.

Network	Number of nodes	Number of links	Degree		Physical connectivity
			Average	Standard Deviation	
NSFNet	16	50	3.1	0.81	0.21
Cost239	11	52	4.7	0.65	0.47
4x4Torus	16	64	4.0	0.00	0.27

Table 5.9: Physical parameters of the topologies used.

Main physical parameters of considered topologies are presented in the Table 5.9, where the average degree is the average number of physical connections available per node. Physical connectivity is defined as the ratio of number of links to the number of links in physically fully connected network of the same size.

5.6.1 Simulation Arrangement

All nodes of the network are joint edge-core nodes as described in Section 4.1 which are connected with packet sources, packet arrivals to the nodes are Poisson with exponentially generated mean packet length of 40 $kbits$ and negative exponential interarrival times. The bursts are assembled using hybrid-based burst assembly mechanism with time threshold of 100 μs and size threshold of 1 $Mbits$. CECS-VF explained in Section 4.2 is used for wavelength scheduling. Burst transmission queues (BTQs) can buffer upto 1000 bursts.

We assume the following:

- The number of wavelengths on each link is 16 with 10 $Gbit/s$ capacity.
- Control packet processing time at a node is 10 μs.
- The offset time of each burst is set according to the number of hops the burst has to traverse. At each hop, offset time is reduced by control pcaket processing time i.e., 10 μs.
- The shortest path obtained in routing is defined as the path with the minimum number of hops.
- The propagation delay per kilometer is 5 μs.
- Full wavelength conversion is available at each node.
- No re-attempt is performed when a connection is blocked.
- The maximum load value for congestion based algorithm (Section 5.4.1) is 0.5.
- The value of weighting factor α, used in weighted link congestion and frequency (WLCF) algorithm (Section 5.4.4) is $\alpha = 0.15$ for NSFNet, $\alpha = 0.95$ for Cost239 and $\alpha = 0.50$ for 4x4Torus topology.
- Link status is updated with the interval of 2 ms using link state protocol (broadcast based approach).
- For exponentially generated moving average, the coefficient x is set to 0.33.

Calculation of offered network load and blocking probabilities are according to the OBS network modelling explained in Section 5.5. Total simulation time is divided into 20 intervals plus a warm-up or transient period and 95% confidence interval used to determine the accuracy of the output. 0.1 million bursts are generated for each simulation interval.

Exponentially Weighted Moving Average

In adaptive routing, when link states are updated to all nodes, all of the nodes start selecting less congested links for burst transmission. This may happen with every update, and the algorithms start oscillating between primary and secondary paths. To avoid the oscillations in the path selection after the update of link status, exponentially weighted moving average (EWMA) is used as shown in the Equation 5.17.

$$S_i = xY_i + (1 - x)S_{i-1} \tag{5.17}$$

Where Y_i is the link status (congestion, load, utilization) at ith update. S_i is the value of EWMA at update i. The coefficient x represents the degree of weighting decrease, a constant smoothing factor between 0 and 1.

Primary and Secondary Path Computations

Primary paths between each source destination pair are computed using Dijkstra algorithm [Dijkstra59]. For the computation of secondary paths, links used in primary path of the considered source destination pair are removed from the topology, and the shortest paths are again calculated using Dijkstra algorithm. Shortest paths are calculated based on the number of hops between each source destination pair. In OBS networks, the more hops a burst has to traverse, the more probability of blocking is there due to the statistical multiplexing. Table 5.10 shows average and standard deviation of number of hops for three considered topologies for calcualated primary and secondary paths.

Path	NSFNet		Cost239		4x4Torus	
	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.
Path I	2.14	1.01	1.42	0.50	2.00	1.00
Path II	3.41	1.22	2.00	0.55	2.50	0.87

Table 5.10: Average and standard deviation of number of hops in primary and secondary paths.

Traffic Scenario

Every node in all simulations is an edge-core joint node (ECJN), connected with traffic generation sources. Simulations are performed using two types of traffic scenarios.

1. **Uniform traffic:** Traffic generated at each source is uniformly distributed to all other nodes. In this case the traffic between each pair of nodes is equal.

2. **Distance dependent traffic:** Traffic between each pair of nodes is inversly proportional to the minimum number of hops between these two nodes. This results in asymetric traffic scenario with less traffic on longer paths.

The distance dependent traffic is generated using multinomial distribution. When a packet is generated at an ingress node, the probability of all the egress nodes is calculated according to the minimum hop counts to reach them using Equation 5.18.

$$p_n = \frac{\frac{1}{h_n}}{\sum_{i=1}^{N} \frac{1}{h_i}} \tag{5.18}$$

Where h_n is the primary path hop count to reach the nth node. N is the total number of hops in the network. The sum of probabilities of all the nodes is 1 as shown in Equation 5.19

$$\sum_{n=1}^{N} p_n = 1 \tag{5.19}$$

These node probabilities are added to make intervals equal in size to the probabilities of nodes, I_n represents the interval for node n as shown in Equation 5.20.

$$I_n = \sum_{i=1}^{N} p_i \qquad n \in \{1, 2, ..., N\} \tag{5.20}$$

Where $I_0 = 0$, $I_1 = p_1$, $I_2 = p_1 + p_2$ and so on upto N. A real number x is generated randomly between 0 and 1. This random number falls in one of the intervals, and classifies the nth destination node for the generated packet at the ingress node using condition ($I_{n-1} > x \geq I_n$).

5.6.2 Results

The primary measure of performance is probability of blocking in OBS networks. Load balancing on alternate paths and better utilization of network resources in terms of increase of throughput is added benefits of adaptive routing algorithms. Following sections show the results of burst blocking probability, load balancing and throughput.

Probability of Blocking

Adaptive routing algorithms improve overall burst blocking probability of the network by chosing alternate paths based on the feedback information about congestion status of the network. Calculations of average offered link load and mean burst blocking probability are based on the formulations given in Section 5.5.

Figure 5.8 plots the mean burst blocking probability of the network as a function average offered link load for NSFNet network with uniform traffic. Blocking probability of selected routing strategies (Section 5.4) are compared with shortest path routing algorithm. The proposed strategy, weighted link congestion and frequency (WLCF) presented in 5.4.4 performs better than other strategies. Figure 5.9 plots the mean burst blocking probability of the network as a function of average offered link load for NSFNet with distance dependent traffic. All of the selected routing strategies show the similar behavior in distance dependent traffic and WLCF performs better than others.

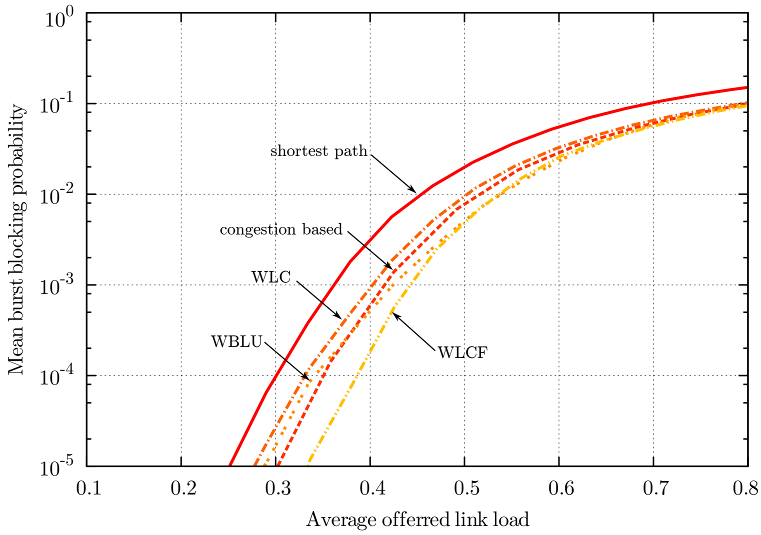


Figure 5.8: Mean blocking probability as a function of average offered link load for NSFNet network with uniform traffic.

Figure 5.10 plots the mean burst blocking probability of the Cost239 network for uniform traffic. It can be observed that WLCF performs better at low loads and higher loads, while WLC strategy performs better in the middle. This is because of the reason that Cost239 network has higher physical connectivity as shown in Table 5.9, and lower average hop distance as shown in Table 5.10, therfore, Cost239 network has less effect of use of frequency in path selection due to higher connectivity. Figure 5.11 shows the same behavior, WLCF strategy performs better at lower and higher loads, and WLC performs better in middle.

Figures 5.12 and 5.13 plot mean blocking probability of the 4x4Torus network with uniform traffic and distance dependent traffic respectively. In both traffic scenarios, WLCF performs much better than other routing

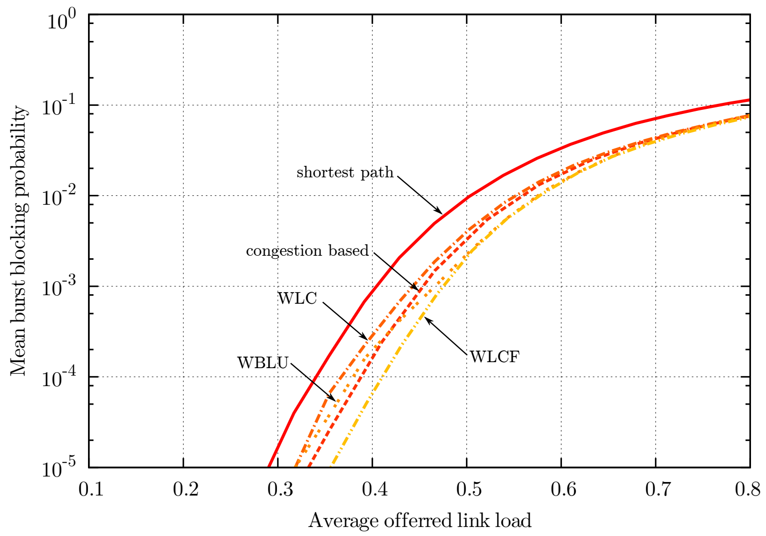


Figure 5.9: Mean blocking probability as a function of average offered link load for NSFNet network with distance dependent traffic.

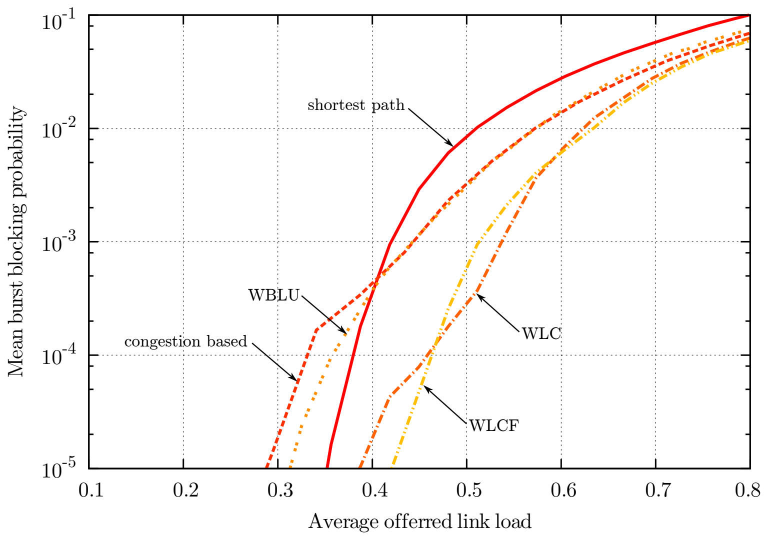


Figure 5.10: Mean blocking probability as a function of average offered link load for Cost239 network with uniform traffic.

strategies. 4x4Torus is has more effect of use of frequency as a path selection parameter, and WLCF algorithm gives advantage to the future connection requests.

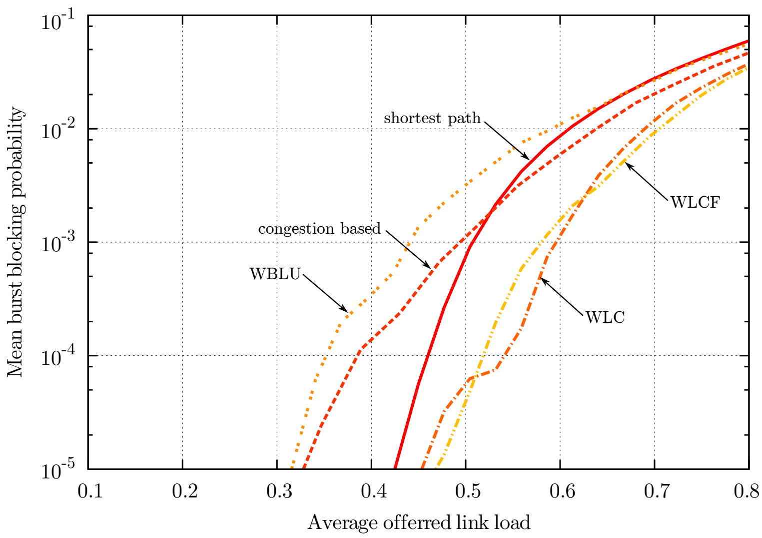


Figure 5.11: Mean blocking probability as a function of average offered link load for Cost239 network with distance dependent traffic.

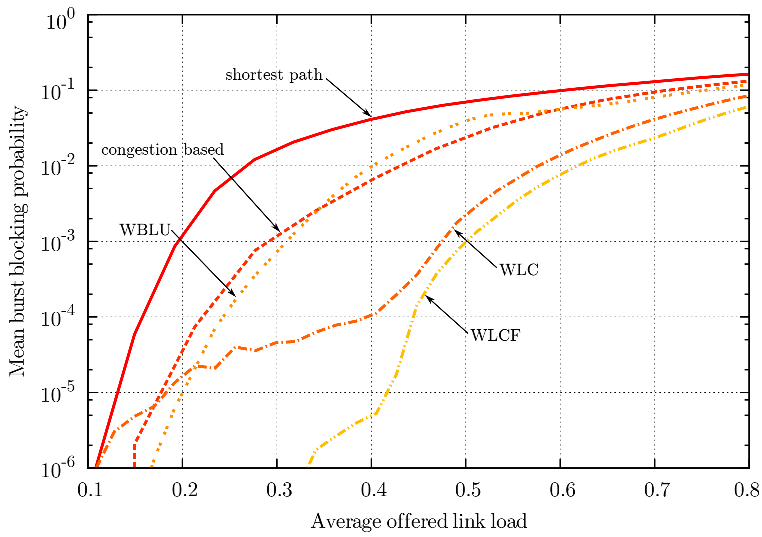


Figure 5.12: Mean blocking probability as a function of average offered link load for 4x4 Torus network with uniform traffic.

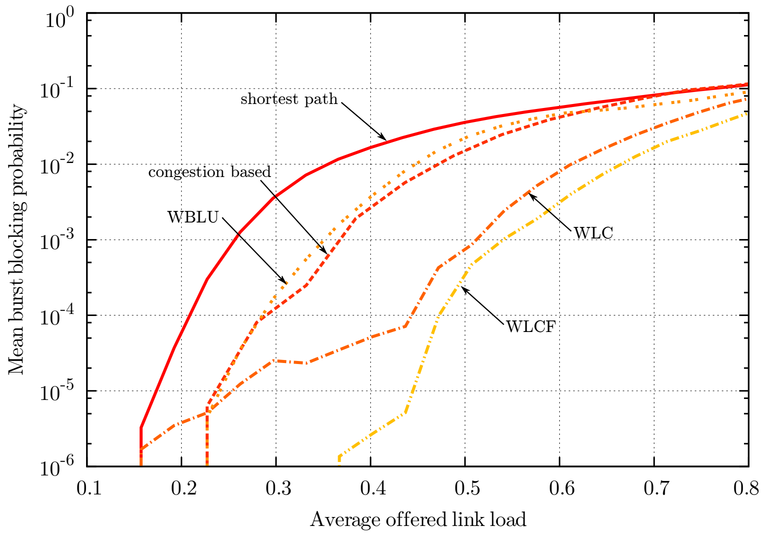


Figure 5.13: Mean blocking probability as a function of average offered link load for 4x4 Torus network with distance dependent traffic.

Load Balancing

Adaptive routing algorithms balance the load by routing the traffic to the less utilized links. Table 5.11 reference the link numbers shown in the load balancing figures and topologies.

Link Number	NSFNet	Cost239	4x4Torus
1	0-1	0-1	0-1
2	0-2	0-2	0-3
3	0-4	0-3	0-4
4	1-0	0-6	0-12
5	1-2	1-0	1-0
6	1-3	1-2	1-2
7	1-8	1-4	1-5
8	2-0	1-7	1-13
9	2-1	2-0	2-1
10	2-6	2-1	2-3
11	3-1	2-3	2-6
12	3-4	2-4	2-14
13	4-0	2-5	3-0
14	4-3	3-0	3-2
15	4-5	3-2	3-7
16	4-11	3-6	3-15
17	5-4	3-7	4-0
18	5-6	3-9	4-5
19	5-7	4-1	4-7
20	6-2	4-2	4-8
21	6-5	4-5	5-1
22	6-10	4-7	5-4
23	6-13	4-10	5-6
24	7-5	5-2	5-9
25	7-8	5-4	6-2
26	8-1	5-6	6-5
27	8-7	5-7	6-7
28	8-8	5-8	6-10
29	9-8	6-0	7-3
30	9-10	6-3	7-4
31	9-12	6-5	7-6
32	9-15	6-8	7-11

Link Number	NSFNet	Cost239	4x4Torus
33	10-6	6-9	8-4
34	10-9	7-1	8-9
35	11-4	7-3	8-11
36	11-12	7-4	8-12
37	11-15	7-5	9-15
38	12-9	7-8	9-8
39	12-11	7-10	9-10
40	12-13	8-5	9-13
41	12-14	8-6	10-6
42	13-6	8-7	10-9
43	13-12	8-9	10-11
44	13-15	8-10	10-14
45	14-12	9-3	11-7
46	14-15	9-6	11-8
47	15-9	9-8	11-10
48	15-11	9-10	11-15
49	15-13	10-4	12-0
50	15-14	10-7	12-8
51	-	10-8	12-13
52	-	10-9	12-15
53	-	-	13-1
54	-	-	13-9
55	-	-	13-12
56	-	-	13-14
57	-	-	14-2
58	-	-	14-10
59	-	-	14-13
60	-	-	14-15
61	-	-	15-3
62	-	-	15-11
63	-	-	15-12
64	-	-	15-14

Table 5.11: Link numbers and actual links on the three topologies.

Figure 5.14 shows the load on every link of the NSFNet network with uniform traffic, while the average link load is approximately 0.5. It can be observed that all of the adaptive routing strategies balance the load in comparison to the shortest path routing. Congestion based routing strategy, described in Section 5.4.1, better balances the load on all the links of the network in comparison to other routing strategies.

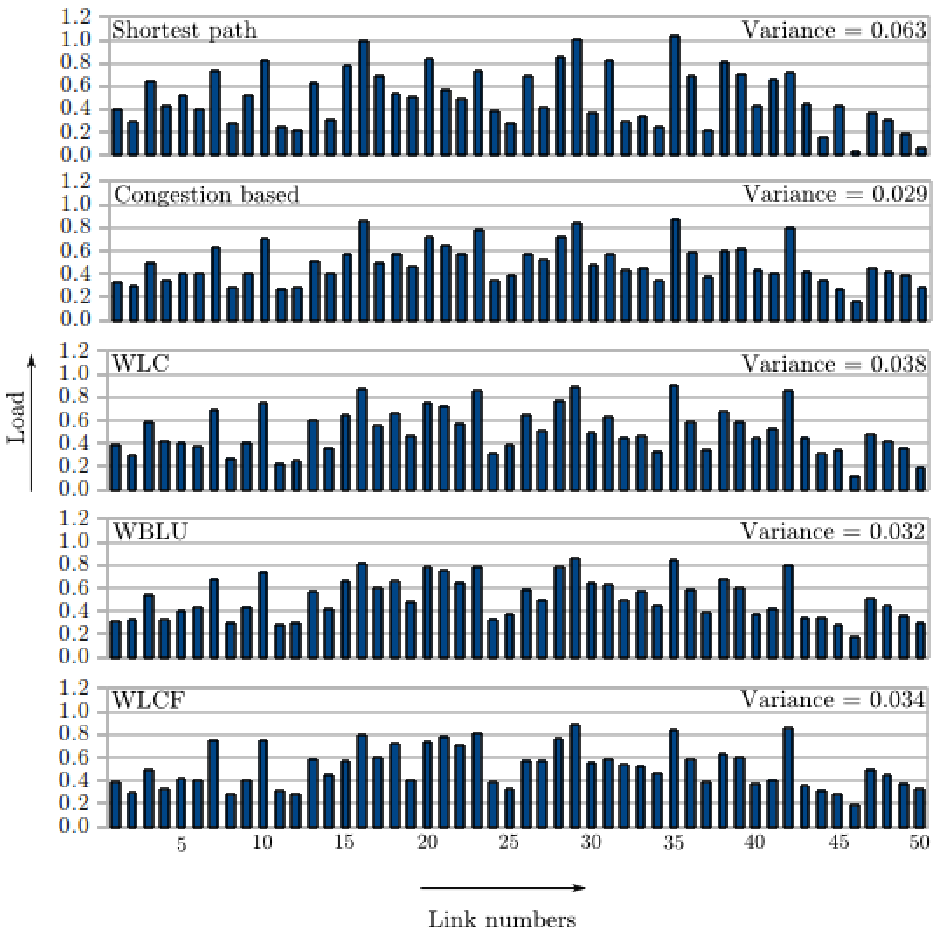


Figure 5.14: Link-wise load for NSFNet network at 0.5 average offered load with uniform traffic.

Figure 5.15 draws the load on every link of NSFNet network with distance dependent traffic. The average link load is approximated to 0.5. We can observe that with distance dependent traffic, WBLU gives better performance, while remaining adaptive routing strategies remain very close to

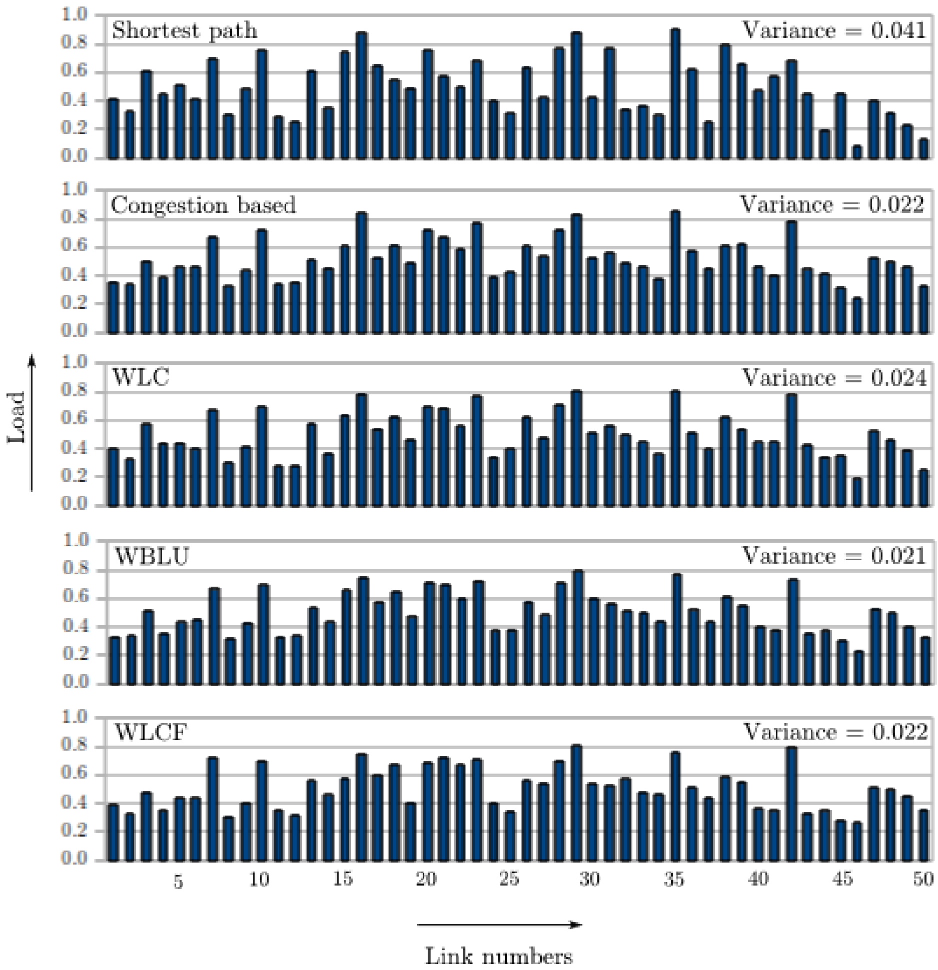


Figure 5.15: Link-wise load for NSFNet network at 0.5 average offered load with distance dependent traffic.

WBLU.

Figures 5.16 and 5.17 display the load on every link in Cost239 network with uniform and distance dependent traffic respectively. The average link load is approximately 0.5. It can be observed from the figures that congestion based routing strategy performs better in Cost239 network in both traffic scenarios, whereas, other adaptive routing strategies perform much better than shortest path.

In case of 4x4Torus network, Figures 5.18 and 5.19 show the link loads.

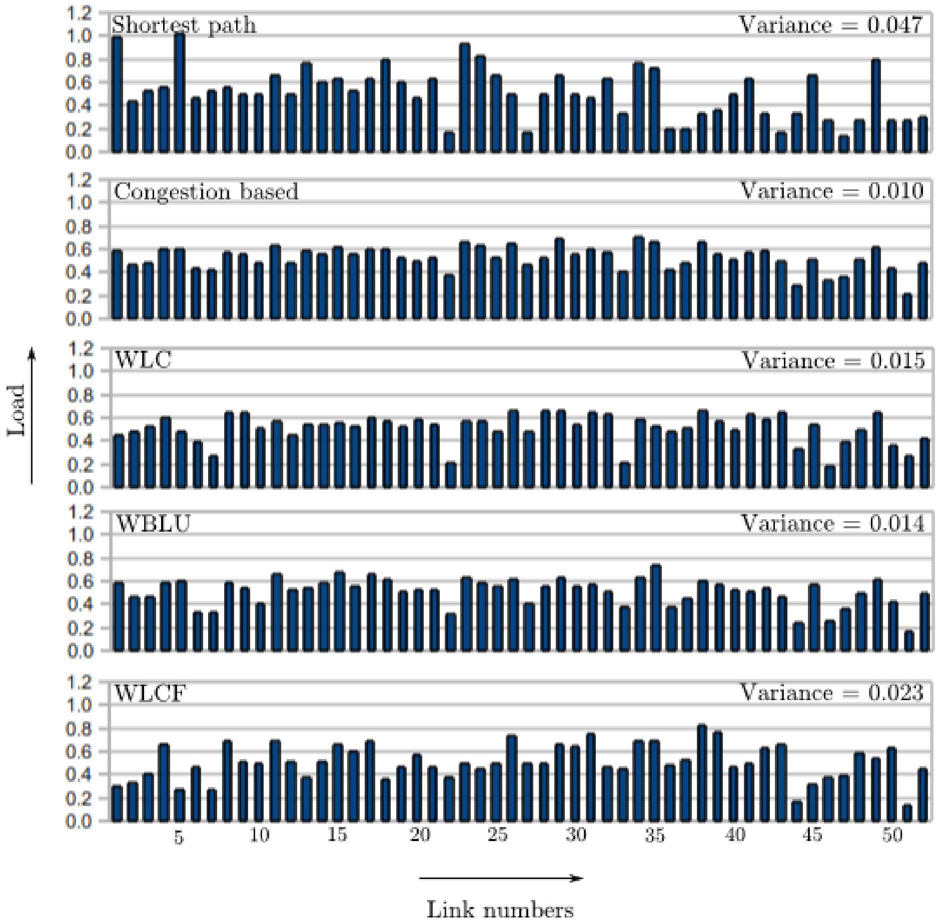


Figure 5.16: Link-wise load for Cost239 network at 0.5 average offered load with uniform traffic.

WLC performs better in terms of load balancing, because 4x4Torus is a symmetrical network and WLC algorithm avoids congested links without going to the longer paths. Congestion based algorithm and WBLU show more load on links 1 to 10, because these links are more frequently used in both primary paths. Whereas, WLCF reduces load on these links because of their high usage in primary paths.

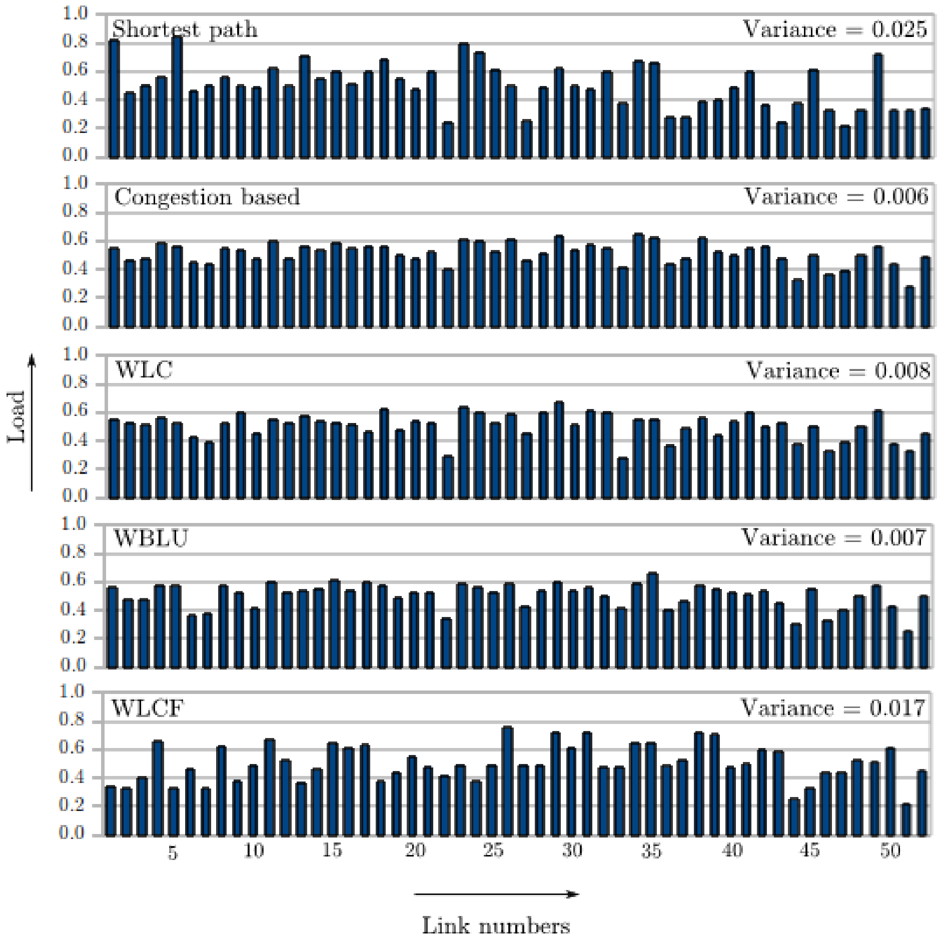


Figure 5.17: Link-wise load for Cost239 network at 0.5 average offered load with distance dependent traffic.

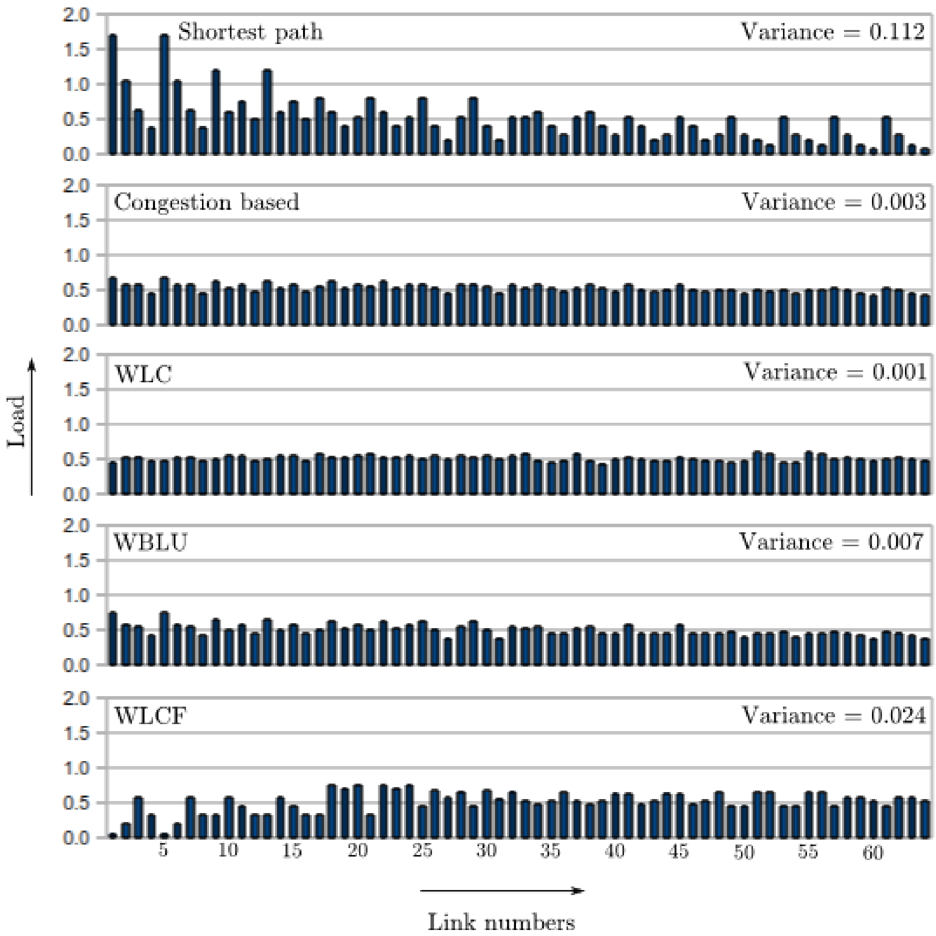


Figure 5.18: Link-wise load for 4x4Torus network at 0.5 average offered load with uniform traffic.

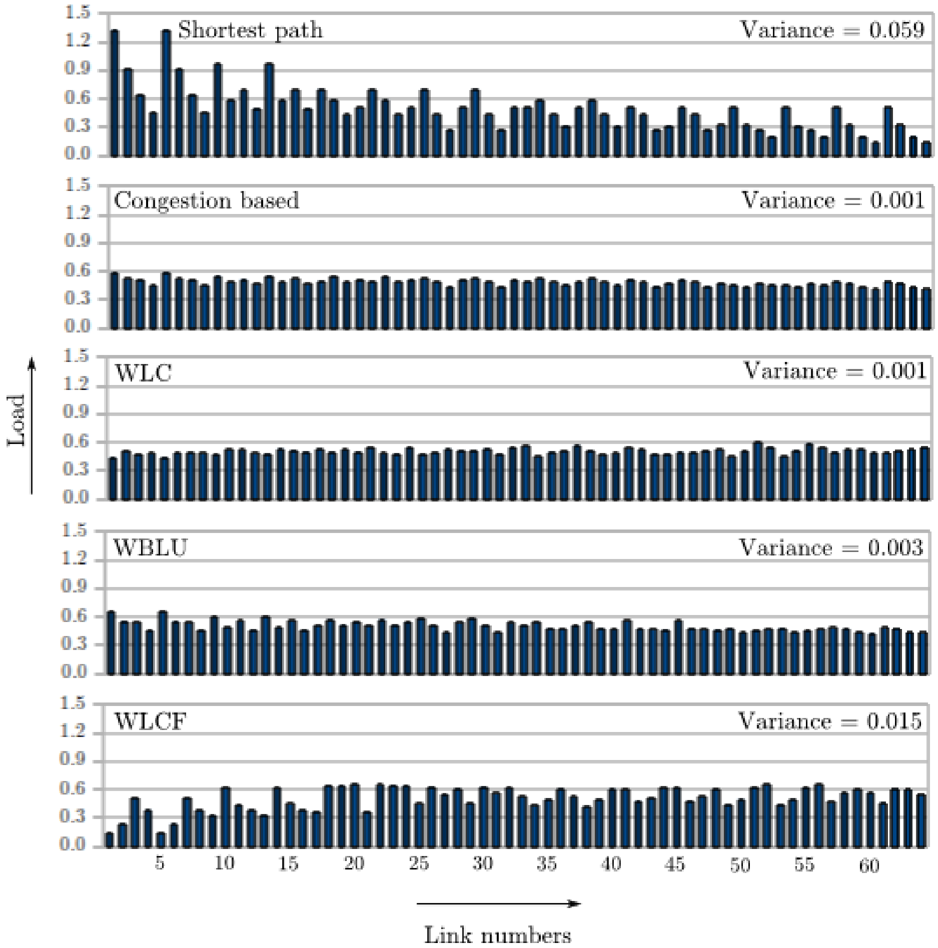


Figure 5.19: Link-wise load for 4x4Torus network at 0.5 average offered load with uniform traffic.

Network Utilization

Adaptive routing algorithms route the traffic on less congested paths, and as a result, network resources are better utilized. Network utilization can be realized by means of network throughput.

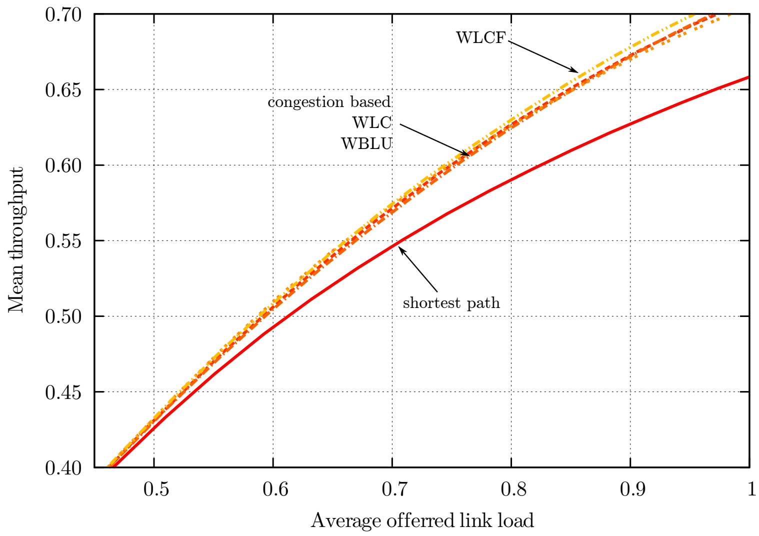


Figure 5.20: Mean network throughput as a function of average offered link load for NSFNet network with uniform traffic.

The calculation of mean network throughput is performed using formulation described in Section 5.5. It is proved by the results that the use of adaptive routing algorithms increases the network throughput. Figure 5.20 and Figure 5.21 plot the mean network throughput as a function of average offered link load for uniform and distance dependent traffic scenarios for NSFNet network. WLCF routing strategy described in Section 5.4.4, performs better in terms of network throughput for both traffic conditions. WLCF reduces the burst blocking probability, and therefore, increases the utilization.

Figure 5.22 and Figure 5.23 show the network throughput for all selected routing strategies for uniform and distance dependent traffic for Cost239 network. All of the selected routing strategies perform better than shortest path routing, and WLCF better utilizes the network and shows high throughput.

Network utilization of 4x4Torus network is shown in Figure 5.24 and Figure 5.25 for uniform and distance dependent traffic. It can be observed that WLCF gives better results in terms of mean network throughput.

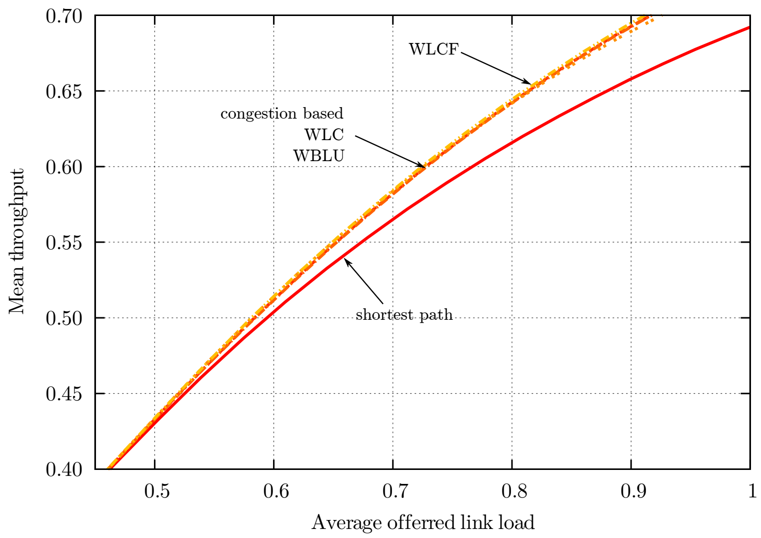


Figure 5.21: Mean network throughput as a function of average offered link load for NSFNet network with distance dependent traffic.

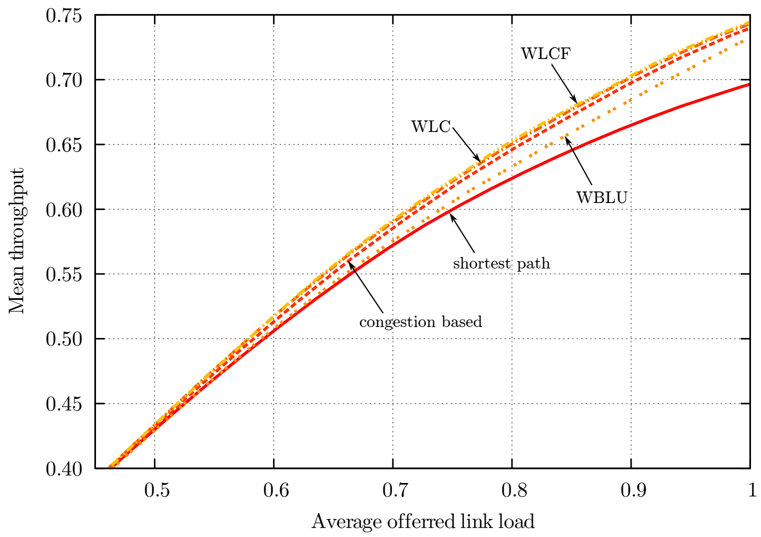


Figure 5.22: Mean network throughput as a function of average offered link load for Cost239 network with uniform traffic.

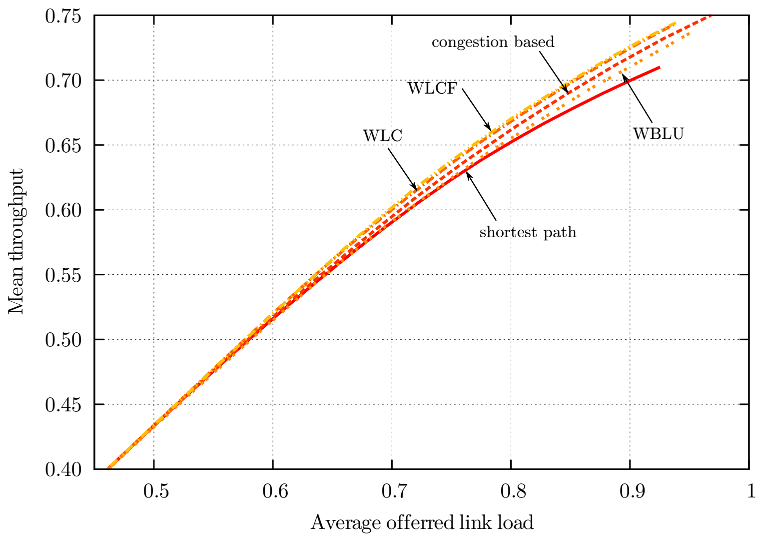


Figure 5.23: Mean network throughput as a function of average offered link load for Cost239 network with distance dependent traffic.

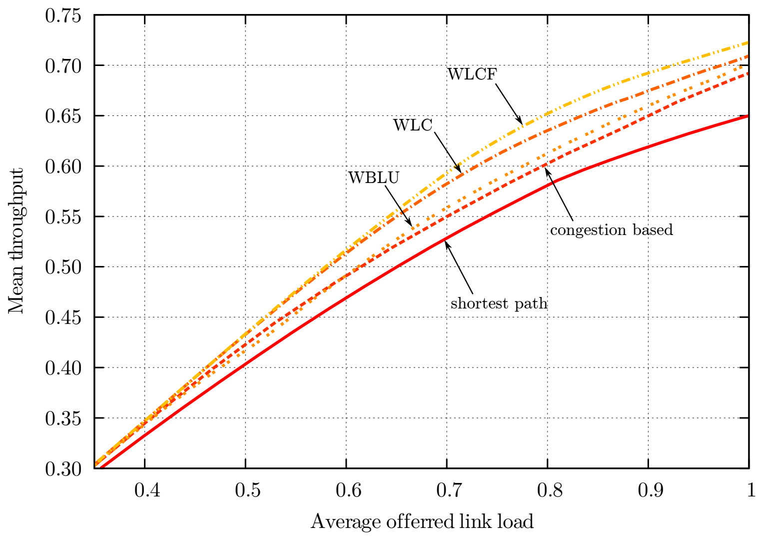


Figure 5.24: Mean network throughput as a function of average offered link load for 4x4 Torus network with uniform traffic.

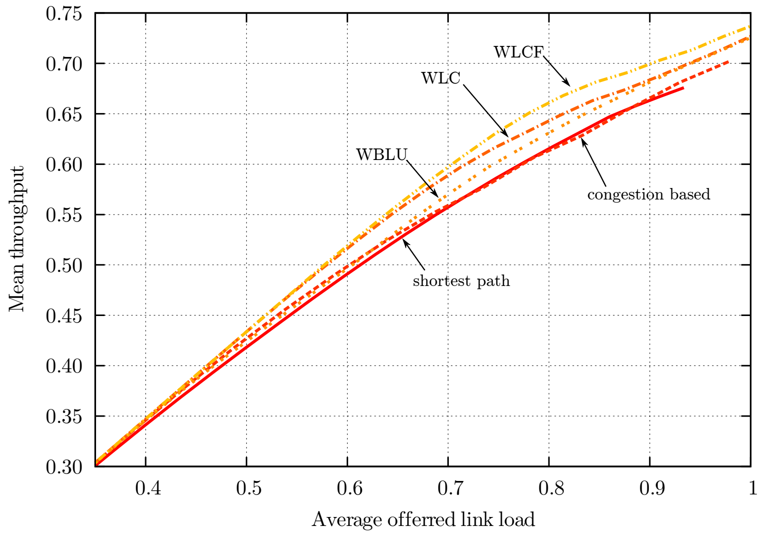


Figure 5.25: Mean network throughput as a function of average offered link load for 4x4 Torus network with distance dependent traffic.

5.7 Summary

In this chapter, the adaptive routing strategies for OBS networks are discussed and simulation results are compared under uniform network operating conditions. Route selection and calculation required in OBS network is considered and different single-path and multi-path routing strategies proposed in literature are studied and classified in Table 5.2 and Table 5.3. Some of the multi-path routing strategies from the literature are selected on the basis of path selection metric. At the end, a new adaptive route selection strategy is proposed. It is proved in the simulation results that the proposed strategy improves burst blocking probability and network throughput.

Chapter 6

Conclusions

In this work, the main focus is the performance analysis of edge-core joint node (ECJN) of optical burst switched networks. The architecture of the joint node and its functionalities are described in Chapter 4. From the results, we can observe that transit and local traffic influence each other. The mixing of both traffic streams is critical part in ECJN. There are several aspects in the mixing of both traffic streams to extend this work. One possible direction is to dedicate certain output channels for each stream and realize the effect with respect to the loss probability and delay. Further possibilities are usage threshold-based schemes with preemptive and non-preemptive appraoches.

Restricted intermediate buffering (RIB) is proposed in Chapter 4 as a contention resolution mechanism. RIB employs that ECJNs may be extended to have the facility to electronically buffer the contending transit bursts, and re-insert them into the network upon availability of resources. It is observed that the intermediate buffering of transit bursts, and prioritizing intermediate burst queue (IBQ) for scheduling causes burst transmission queue (BTQ) to overflow. Implementation of intermediate buffering needs certain criteria to schedule buffered local and transit bursts.

OBS is supposed to be a backbone network, and dropped packets are retransmitted from upper network layers. The study of intermediate buffering may be extended to compare the results with retransmission of data.

Adaptive routing strategies for OBS networks are discussed in detail in Chapter 5. Adaptive multipath routing algorithms proposed in literature are categorized in Table 5.3. It is observed that criterion for path selection is based on link loads, link utilization and link congestion. One routing strategy from each path selection criterion is selected and simulated under uniform network parameters. A new route selection strategy is proposed by taking into account the link congestion and frequency of the links' usage in primary paths. Results show that the proposed strategy performs better in terms of burst loss rate and network throughput. The obtained results also show that the adaptive routing strategies balance the load on network links in comparison to the shortest path.

The path selection in adaptive routing is based on the current status of the links in the available paths. The status of the links is updated periodically using probe based or broadcast based approach. The interval of status message update is based on several parameters, too frequent status message updates may cause high signaling overhead, and less frequent status message update may cause outdated infomation for the decision of path selection. The study of adaptive routing strategies may be extended by realizing the effect of changes in status message update timing.

There are possibilities of extending the study of adaptive routing strategies in route calculation phase. Some effects of link congestion can be addressed in route calculation phase, and better results may be achieved in path selection phase.

List of Figures

List of Tables

Bibliography

[Agusti04] A. Agusti and C. Cervello-Pastor. A new contentionless dynamic routing protocol for OBS using wavelength occupation knowledge. In *12th IEEE Mediterranean Electrotechnical Conference. MELECON*, 2004.

[Argos07] C.G. Argos, O.G. de Dios, and J. Aracil. Adaptive Multi-Path Routing for OBS Networks. In *9th International Conference on Transparent Optical Networks. ICTON '07.*, 2007.

[Barradas09] Alvaro L. Barradas and Maria do Carmo R. Medeiros. Edge-Node Deployed Routing Strategies for Load Balancing in Optical Burst Switched Networks. *ETRI Journal*, vol. 31, no. 1, pp. 31–41, 2009.

[Barradas10] Alvaro L. Barradas and Maria do Carmo R. Medeiros. Pre-planned optical burst switched routing strategies considering the streamline effect. *Photonic Network Communications*, vol. 19, no. 2, pp. 161–169, April 2010.

[Bikram10] R.R.C. Bikram, N. Charbonneau, and V.M. Vokkarane. Coordinated Multi-Layer Loss Recovery in TCP over Optical Burst-Switched (OBS) Networks. In *IEEE International Conference on Communications (ICC).*, pages 1–5, May 2010.

[Boudriga03] Noureddine Boudriga. Optical burst switching protocols for supporting QoS and adaptive routing. *Computer Communications: Recent Advances in Computer Communications networking*, vol. 26, no. 15, pp. 1804–1812, 2003.

[Callegati06] F. Callegati, W. Cerroni, and C. Raffaelli. Impact of Optical Packet Loss and Reordering on TCP Performance. In *Global Telecommunications Conference. IEEE GLOBECOM '06.*, 2006.

[Chen06] Qian Chen, Gurusamy Mohan, and Kee Chaing Chua. Offline Route Optimization Considering Streamline Effect in Optical

Burst Switching Networks. In *IEEE International Conference on Communications. ICC '06.*, 2006.

[Chen08] Qian Chen, Gurusamy Mohan, and Kee Chaing Chua. Route optimization in optical burst switched networks considering the streamline effect. *Computer Networks: Challenges and Opportunities in Advanced Optical Networking*, vol. 52, pp. 2033–2044, 2008.

[Chiaroni95] D. Chiaroni. A novel photonic architecture for high capacity atm switch applications. In *Proceedings of the Photonics in Switching conference, Salt Lake City, UT.*, April 95.

[Chlamtac92] I. Chlamtac, A. Ganz, and G. Karmi. Lightpath communications: an approach to high bandwidth optical WAN's. *IEEE Transactions on Communications*, vol. 40, no. 7, pp. 1171–1182, Jul 1992.

[Coulibaly11] Yahaya Coulibaly, Muhammad Abd Latiff Shafie, and Abu Bakar Mohammad. A review of routing strategies for optical burst switched networks. *International Journal of Communication Systems*, , 2011.

[Diaz09] Mariana Escolar Diaz and Xavier Hesselbach. Routing with prioritization based on statistics in OBS networks. In *11th International Conference on Transparent Optical Networks. ICTON '09.*, pages 1–4, 2009.

[Dijkstra59] E.W. Dijkstra. A note on two problems in connection with graphs. *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[Du06] Yu Du, Tao Pu, Hanyi Zhang, and Yili Guo. Adaptive load balancing routing algorithm for optical burst-switching networks. In *Optical Fiber Communication Conference, National Fiber Optic Engineers Conference, OFC 2006*, 2006.

[Eramo03] V. Eramo, M. Listanti, and P. Pacifici. A comparison study on the wavelength converters number needed in synchronous and asynchronous all-optical switching architectures. *IEEE Journal of Lightwave Technology*, vol. 21, no. 2, pp. 340–355, 2003.

[Ganguly04] S. Ganguly, S. Bhatnagar, R. Izmailov, and Chunming Qiao. Multi-path adaptive optical burst forwarding. In *Workshop on High Performance Switching and Routing, HPSR.*, 2004.

[Gao08] Xingbo Gao and M.A. Bassiouni. Fairness-Improving Adaptive Routing in Optical Burst Switching Mesh Networks. In *International Conference on Communications. ICC '08. IEEE*, 2008.

[Gao09] Xingbo Gao, Mostafa A. Bassiouni, and Guifang Li. Addressing conversion cascading constraint in OBS networks through proactive routing. *Photonic Network Communications*, vol. 18, no. 1, pp. 90–104, 2009.

[Garcia07] Nuno M. Garcia, Prezemyslaw Lenkiewicz, Mrio M. Freire, and Paulo P. Monteiro. On the Performance of Shortest Path Routing Algorithms for Modeling and Simulation of Static Source Routed Networks – an Extension to the Dijkstra Algorithm. In *Second International Conference on Systems and Networks Communications. ICSNC 2007.*, 2007.

[Garcia08] Nuno Manuel Garcia. *Architectures and Algorithms for IPv4/IPv6-Compliant Optical Burst Switching Networks.* PhD thesis, Department of Computer Science, University of Beira Interior, Covilh, Portugal., 2008.

[Ge00] A. Ge, F. Callegati, and L.S. Tamil. On optical burst switching and self-similar traffic. *IEEE Communications Letters*, vol. 4, no. 3, pp. 98–100, March 2000.

[Gunreben11] Sebastian Gunreben. An optical burst reordering model for time-based and random selection assembly strategies. *Performance Evaluation*, vol. 68, no. 3, pp. 237–255, March 2011.

[Hayat11a] Muhammad Faisal Hayat and Tomasz Hoyski. Modelling of OBS hybrid burst assembly using transform based approach. In *1st European Teletraffic Seminar (ETS)*, 2011.

[Hayat11b] Muhammad Faisal Hayat, Farrukh Zeeshan Khan, and H.R. van As. Performance model for an OBS node with a shared wavelength converter pool and an FDL buffer per link. In *15th International Conference on Optical Network Design and Modeling (ONDM).*, pages 1–6, February 2011.

[Hayat12] Muhammad Faisal Hayat, Tomasz Hoyski, and Farrukh Zeeshan Khan. Performance study of OBS combined node via mixed loss-delay queueing systems. In *Submitted to 24th International Teletraffic Congress (ITC 24), Krakow, Poland*, 2012.

[Hendling04] Karl Hendling. *Efficient On-Line Routing Strategies of Bandwidth Guaranteed Paths with Traffic Engineering.* PhD thesis, Institute of Broadband Communication, Technical University, Vienna, Austria, 2004.

[Hirota06] Y. Hirota, H. Tode, and K. Murakami. Cooperation method considering wavelength assignment and routing problem in optical burst switched networks. In *National Fiber Optic Engineers Conference (NFOEC)*, 2006.

[Huang05a] X. Huang and V.M. Vokkarane. Burst cloning: a proactive scheme to reduce data loss in optical burst-switched networks. In *IEEE International Conference on Communications. ICC 2005.*, volume 3, pages 1673 – 1677, May 2005.

[Huang05b] Y. Huang, J.P. Heritage, and B. Mukherjee. Dynamic routing with preplanned congestion avoidance for survivable optical burst-switched (OBS) networks. In *Optical Fiber Communication Conference.*, 2005.

[Hudek95] G. Hudek and D. Muder. Signaling analysis of a multi-switch all-optical network. In *IEEE International Conference on Communications (ICC).*, pages 1206–1210, 1995.

[Hyyti04] Esa Hyyti and Laura Nieminen. Linear program formulation for routing problem in OBS networks. In *Ninth International Symposium on Computers and Communications, ISCC.*, 2004.

[Ishii05] D. Ishii, N. Yamanaka, and I. Sasae. A self-learning route selection scheme using multi-path searching packets in an OBS network. In *Workshop on High Performance Switching and Routing, HPSR.*, pages 346 – 350, 2005.

[Jue05] Jason P. Jue and Vinod M. Vokkarane. *Optical Burst Switched Networks.* Springer, 2005.

[Kaheel03] A. Kaheel and H. Alnuweiri. A Strict Priority Scheme for Quality-of-Service provisioning in Optical Burst Switching Networks. In *8th IEEE International Symposium on Computers and Communications (ISCC 2003).*, volume 1, pages 16–21, 2003.

[Kim02] Sungchang Kim, Namook Kim, and Minho Kang. Contention resoultion for optical burst switching networks using alternative routing. In *IEEE International Conference on Communications, ICC 2002.*, volume 5, pages 2678 – 2681, 2002.

[Klinkowski05] M. Klinkowski, D. Careglio, S. Spadaro, and J.S. Pareta. Impact of Burst Length Differentiation on QoS performance in OBS networks. In *7th International Conference on Transparent Optical Networks (ICTON 2005)*, volume 1, pages 91–94, Jul 2005.

[Klinkowski07a] M. Klinkowski, M. Pioro, D. Careglio, M. Marciniak, and J. Sole-Pareta. Non-linear optimization for multi-path source routing in obs networks. *Communications Letters, IEEE*, vol. 11, no. 12, pp. 1016 – 1018, 2007.

[Klinkowski07b] Miroslaw Klinkowski. *Offset Time-Emulated Architecture for Optical Burst Switching - Modelling and Performance Evaluation.* PhD thesis, Departament d'Arquitectura de Computadors, Universitat Politecnica de Catalunya, 2007.

[Klinkowski10] Mirosaw Klinkowski, Joo Pedro, Davide Careglio, Micha Piro, Joo Pires, Paulo Monteiro, and Josep Sol-Pareta. An overview of routing methods in optical burst switching networks. *Optical Switching and Networking*, vol. 7, no. 2, pp. 41–53, January 2010.

[Lee05] S. Lee, L. Y. Kim, J. S. Song, D. W. Griffith, and K. Sriram. Dynamic Deflection Routing with Virtual Wavelength Assignment in Optical-Burst Switched Networks. *Photonic Network Communications*, vol. 9, no. 3, pp. 347–356, May 2005.

[Lee06] SeoungYoung Lee, InYong Hwang, and HongShik Park. A new burst scheduling algorithm for edge/core node combined optical burst switched networks. In *NETWORKING'06 Proceedings of the 5th international IFIP-TC6 conference on Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems*, 2006.

[Levesque09] Martin Levesque and Halima Elbiaze. Graphical probabilistic routing model for OBS networks with realistic traffic scenario. In *Proceedings of the 28th IEEE conference on Global telecommunications, GLOBECOM'09 .*, 2009.

[Lewis72] Peter A. W. Lewis. *Stochastic point processes: statistical analysis, theory, and applications.* Wiley Inter-science Publication, 1972.

[Li04] J. Li, G. Mohan, and K.C. Chua. Dynamic load balancing in IP-over-WDM optical burst switching networks. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 47, no. 3, pp. 393408, 2004.

[Li06] Ji Li and K.L. Yeung. Burst cloning with load balancing. In *Optical Fiber Communication Conference, OFC.*, 2006.

[Li07] Hailong Li and Ian Li-Jin Thng. Edge node buffer usage in optical burst switching networks. *PHOTONIC NETWORK COMMUNICATIONS*, vol. 13, no. 1, pp. 31–51, 2007.

[Lu06] Jia Lu, Yong Liu, M. Gurusamy, and K.C. Chua. Gradient projection based multi-path traffic routing in optical burst switching networks. In *Workshop on High Performance Switching and Routing.*, 2006.

[Ou05] Liang Ou, Xiansi Tan, Huaxiong Yao, and Wenqing Cheng. A Study on Dynamic Load Balanced Routing Techniques in Time-Slotted Optical Burst Switched Networks. *Proceedings of the Third international conference on Networking and Mobile Computing*, vol. 3619, pp. 722–731, 2005.

[Pedro07a] Joao Pedro, Paulo Monteiro, and Joao Pires. Efficient Optical Burst-Switched networks using only Fiber Delay Line buffers for contention resolution. In *Fourth International Conference on Broadband Communications, Networks and Systems*, pages 2–11, Sep 2007.

[Pedro07b] Joo Pedro, Jorge Castro, Paulo Monteiro, and Joo Pires. Efficient Multi-Path Routing for Optical Burst-Switched Networks. In *Proceedings of Conference on Telecommunications, ConfTele, Peniche, Portugal*, 2007.

[Pedrola10] Oscar Pedrola, Miroslaw Klinkowski, Davide Careglio, Josep Sol-Pareta, Sbastien Rumley, and Christian Gaumier. JAVOBS: A Flexible Simulator for OBS Network Architectures. *Journal of Networks*, vol. 5, no. 2, pp. 256–264, Feb 2010.

[Qiao99] C. Qiao and M. Yoo. Optical burst switching (OBS) - a new paradigm for an optical Internet. *Journal of High Speed Networks*, vol. 8, no. 1, pp. 69–84, 1999.

[Rodrigues04] Joel J. P. C. Rodrigues, Nuno M. Garcia, Mrio M. Freire, and Pascal Lorenz. Object-Oriented Modeling and Simulation of Optical Burst Switching Networ. In *IEEE Global Communications Conference Workshops (GLOBECOM 2004)*, 2004.

[Rosberg03] Z. Rosberg, Hai Le Vu, M. Zukerman, and J. White. Blocking probabilities of optical burst switching networks based on reduced load fixed point approximations. In *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. INFOCOM 2003.*, 2003.

[Sarwar08] S. Sarwar, S. Aleksic, and K. Aziz. Optical burst switched (OBS) systems with segmentation-based dropping. *E & I, ELEKTROTECHNIK UND INFORMATIONSTECHNIK*, vol. 125, no. 7-8, pp. 296–300, 2008.

[Schuringa04] J.H. Schuringa. *Mobile Routing Agents deployed in Resilient Packet Networks*. PhD thesis, Vienna University of Technology, Vienna, Austria, 2004.

[Tan04] C. W. Tan, M. Gurusamy, and J. C. S. Lui. Achieving Proportional Loss Differentiation using Probabilistic Preemptive Burst

Segmentation in Optical Burst Switching WDM Networks. In *IEEE Global Telecommunications Conference (GLOBECOM 2004)*, volume 3, pages 1754–1758, 2004.

[Teng05] Jing Teng and George N. Rouskas. Routing path optimization in optical burst switched networks. In *Conference on Optical Network Design and Modeling.*, pages 1–10, 2005.

[Thodime03] G.R.V. Thodime, V.M. Vokkarane, and J.P. Jue. Dynamic congestion-based load balanced routing in optical burst-switched networks. In *IEEE Global Telecommunications Conference (GLOBECOM '03)*, 2003.

[Triay10] J. Triay and C. Cervello-Pastor. An ant-based algorithm for distributed routing and wavelength assignment in dynamic optical networks. *IEEE Journal on Selected Areas in Communications.*, vol. 28, no. 4, pp. 542–552, 2010.

[Turner99] Jonathan S. Turner. Terabit Burst Switching. *Journal of High Speed Networks*, vol. 8, no. 1, pp. 3–16, March 1999.

[Ueda05] Masayuki Ueda, Takuji Tachibana, and Shoji Kasahara. Last-hop preemptive scheme based on the number of hops for optical burst switching networks. *Journal of Optical Networking*, vol. 4, no. 10, pp. 648–660, 2005.

[Vokkarane02a] Vinod M. Vokkarane. Burst segmentation: an approach for reducing packet loss in optical burst switched networks. In *IEEE International Conference on Communications, ICC 2002.*, volume 5, pages 2673 – 2677, May 2002.

[Vokkarane02b] Vinod M. Vokkarane, Karthik Haridoss, and Jason P. Jue. Threshold-based burst assembly policies for QoS support in optical burst-switched networks. In *Optical Networking and Communications, OptiComm 2002*, July 2002.

[Vokkarane02c] Vinod M. Vokkarane, Qiong Zhang, Jason P. Jue, and Biao Chen. Generalized burst assembly and scheduling techniques for QoS support in optical burst-switched networks. In *IEEE Global Telecommunications Conference, GLOBECOM '02.*, Nov 2002.

[Vokkarane03] Vinod Vokkarane, Guru Thodime, Venkata Challagulla, and Jason Jue. Channel Scheduling Algorithms using Burst Segmentation and FDLs for Optical Burst-Switched Networks. In *IEEE International Conference on Communications (ICC).*, 2003.

[Wallentin10] Lukas Wallentin, Marco Happenhofer, Christoph Egger, and Joachim Fabini. XML meets Simulation: Concepts and Architecture of the IBKSim Network Simulator. *Eurosim - Simulation News Europe*, vol. 1, pp. 16–20, 2010.

[Wang00] Xi Wang, Hiroyuki Morikawa, and Tomonori Aoyama. Deflection routing protocol for burst switching WDM mesh networks. *Terabit Optical Networking: Architecture, Control, and Management Issues*, vol. 4213, pp. 242–252, 2000.

[Widjaja95] I. Widjaja. Performance analysis of burst admission-control protocols. *IEEE Proceeding of Communications*, vol. 142, pp. 7–14, 1995.

[Xiong99] Yijun Xiong, Marc Vandenhoute, and Hakki C. Cankaya. Design and analysis of optical burst-switched networks . *All-Optical Networking: Architecture, Control, and Management Issues*, vol. 3843, pp. 112–119, Aug 1999.

[Xiong00] Yijun 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, October 2000.

[Xu03] J. Xu, Chunming Qiao, J. Li, and G. Yu. Efficient Channel Scheduling Algorithms in Optical Burst Switched Networks. *IEEE Infocomm*, vol. 3, pp. 2268–2278, 2003.

[Yang06] Li Yang and George N. Rouskas. Adaptive Path Selection in OBS Networks. *Journal of Lightwave Technology*, vol. 24, no. 8, pp. 3002–3011, 2006.

[Yoo97] M. Yoo and Chunming Qiao. Just-enough-time (JET): a high speed protocol for bursty traffic in optical networks. *IEEE Photonics Society (IPS).*, vol. 3, no. 11, pp. 26–27, Aug 1997.

[Yoo00] M. Yoo, Chunming Qiao, and S. Dixit. QoS Performance of Optical Burst Switching in IP-over-WDM Networks. *IEEE Journal of Selected Areas in Communications.*, vol. 18, pp. 2062–2071, 2000.

[Yu02] Xiang Yu, Yang Chen, and Chunming Qiao. Study of traffic statistics of assembled burst traffic in optical burst-switched networks. In *Optical Networking and Communications, OptiComm 2002*, 2002.

[Yuan11] Chi Yuan, Zhenrong Zhang, Zhengbin Li, Yongqi He, and Anshi Xu. A unified study of burst assembly in optical burst switching networks. *Photonic Network Communications*, vol. 21, no. 3, pp. 228–237, June 2011.

[Zhang04a] Jing Zhang, Hae-Joung Lee, Shuang Wang, Xue Qiu, Keyao Zhu, Yurong Huang, Debasish Datta, Young-Chon Kim, and Biswanath Mukherjee. Explicit Routing for Traffic Engineering in Labeled Optical Burst-Switched WDM Networks. *Proceedings of Internation Conference on Computational Science*, vol. 3038, pp. 1209–1216, 2004.

[Zhang04b] Jing Zhang, Shuang Wang, Keyao Zhu, D. Datta, Young-Chon Kim, and B. Mukherjee. Pre-planned global rerouting for fault management in labeled optical burst-switched WDM networks. In *IEEE Global Telecommunications Conference, GLOBECOM '04.*, 2004.