Durban University of Technology

Faculty of Engineering and the Built Environment

Department of Electronic Engineering

Contention and Congestion Minimization in OBS Networks

Ndadzibaya Masimba Gomba [21451055]

This Research is submitted in partial fulfilment for the Masters in Engineering Degree - Electrical Engineering.

Supervisor: Prof B. Nleya Date.........................

Durban, 2016

Declaration

I declare that this work is my own. Where collaboration with other people has taken place, or material entered by other researchers is included, the parties and/or martial are indicated in the acknowledgement or references as appropriate.

This work is being submitted for the degree of Master of Engineering (Electrical) in the Department of Electronic Engineering. It has not been submitted to any other university for any other degree or examination.

Masimba Gomba Ndadzibaya Date

Acknowledgement

I would also like to thank my family. To my colleagues at SA Home Loans and CTI Educational Group, thank you very much for all the support and understanding. Your words of encouragement and confidence in me really helped a lot.

Abstract

All-optical networks (AON) based optical burst switching (OBS) promise to be the ultimate backbone network technology solution for next generation(NG) as well as future generation (FG)networks because of their relatively higher resources utilization, great flexibility at lower cost and potential massive bandwidth capacities both at transmission and switching levels. By design, buffering is not provisioned in interior nodes. End users exchange data with one another through end-to-end light channels, called lightpaths in which wavelength continuity is maintained. In practice, their establishing, in a cost effective manner remains an inescapable challenge. The routing and wavelength assignment (RWA) problem entices successful establishment of a physical route for each lightpath connection request, assigning a wavelength to each route and at the same time ensuring end to end continuity, subject to the limited number of wavelengths. The wavelengths must be assigned such that no lightpaths can share the same wavelength simultaneously on a given fibre, or else contentions may occur. Some data bursts may be discarded whenever contention occurs among multiple bursts that arrive simultaneously at any interior node using the same wavelength and are intended for the same output port. Because of the buffer-less nature of OBS networks, contention/congestion in the core network can quickly lead to degradation in overall network performance at moderate to high traffic levels due to heavy burst loses. In this dissertation we propose and evaluate a congestion management approach we refer to as “enhanced congestion management" which gears towards rendering and guaranteeing a consistent QoS as well as rational and fair use of available network links. Simulation results show that the scheme can effectively minimize both contention and congestion and at the same time improving both throughput and effective utilization under moderate to high network traffic loads.

Table of Content

List of Figures

List of Tables

List of acronyms

BCP	-	Burst Control Packet
CP	-	Control Packet
BSR	-	Burst Size Ratio
FDLs	-	Fibre Delay Lines
FSN	-	Forwarding and Switching Nodes
JET	-	Just-Enough-Time
JIT	-	Just-In-Time
LP	-	Low Priority
NGN	-	Next Generation Networks
O/E/O	-	Optical-to-Electronic-to-Optical
OBS	-	Optical Burst Switching
OCS	-	Optical Circuit Switching
OLS	-	Optical Label Switching
OPS	-	Optical Packet Switching
OXC	-	Optical Cross Connection
PDF	-	The probability density function
QoS	-	Quality of Service
QoX	-	Quality of Experience
WDM	-	Wavelength Division Multiplexing

Chapter 1. Introduction

The ever-continuing advancements in communications and networking technologies and capabilities have surpassed unprecedented heights. Various technologies and approaches continue to provide pervasive network connectivity not only to the home users and workplace but also to the countryside where no wired infrastructure can reach because of the unfriendly nature of the terrain. In addition, recent advances in optical networking have vastly improved transmission and switching capacities of core and access telecommunication networks. Ultimately these advances have impacted significantly on our everyday lives and hence many new challenges and opportunities are emerging.

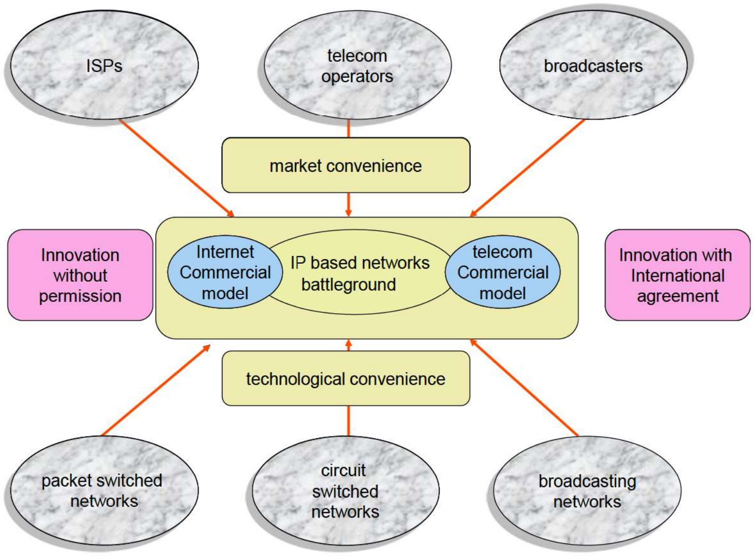

Figure 1-1: NGN Convergence.

Of particular importance is the emergence of the Next Generation Networking (NGN) platform in which all the various existing access technologies share the same core network. As a consequence, the support for, participating network devices, heterogeneity in services as well as transportation medium is driving the continuous evolving of the NGN [1].

The traffic in NGN is supported primarily on personal devices as well as relatively unconstrained desktop computers , all of which may be constrained by power, bandwidth availability as well as processing capability. The propagation medium is increasing on-uniformly varying between wireless and wired environments. The traffic is surging in volumes and mostly exhibiting stochastic /non-stationary statistical behaviour. The ever increasing heterogeneity, results in expectations from the network regarding the QoS are becoming more complex and demanding. The real-time nature of significant traffic such as video, multimedia and voice is poses high and complex QoS requirements. Various traffic classes variably contribute to the complex aggregated traffic patterns and as such the QoS requirements will vary considerably from time to time.

1.1 Transmission and Switching in OBS networks

The ever-surging traffic levels in the envisaged amalgamated universal network brought about the desire to design an appropriate switching technology for an optical backbone network that would meet both the capacity as well as desired QoS demands. Of the three primary candidates; optical burst switching (OBS), optical circuit switching (OCS) and optical packet switching (OPC), OBS eventually emerged as a perspective and practical switching solution. It's a switching paradigm that is conceptualized and ultimately designed inheriting and amalgamating the best features of both OCS and OPS. OBS thus provisions better bandwidth efficiency, and at the same time it eliminates the need for buffering in the core network as well as optical –to- electronic conversions and vice-versa of the data bursts at the core (switching) nodes hence effectively narrowing the gap between transmissions and switching speeds in FGNs. At the physical layer (transmission level), data packets tapped from ingress nodes are aggregated and ultimately assembled into optical huge packets generally referred to as bursts. An associated burst control header packet (CHP) is transmitted for each burst via a reserved control channel (CC) and received at the next node with a carefully set small relative offset-time ahead of the arrival time of the actual data burst. The offset timing generally allows for sufficient time in electronic *processing* of the CHP by dedicated controllers at an intermediate nodes thus creating time for reserving a wavelength in its output port/ link, normally for a duration equaling that of the incoming burst, as well as to reconfigure the switching fabric. The burst will thus shortly fly-by and immediately afterwards the reserved resources are released and made available for any other new connection demands. Hence this removes the need for buffering at interior nodes which normally could further intricate network engineering design as well as operational costs and latencies. Furthermore, this momentary usage of these resources promotes enhanced overall utilization of network resources as well as improved adaptation to bursty traffic sources in comparison to optical the traditional circuit-switching paradigms and networks. Since OBS networks utilizes limited buffering it thus will always be susceptible congestion.

The presence of overloaded links may aid to a serious aggravation of the network's throughput [2]. The congestion can partly be decreased by either careful network re-engineering or by proper proactive routing approaches in the core network. The dimensioning approach fixes the node/link resources according to the matrix of actual traffic load demands and after careful optimization, it needs only either a simplified shortest path algorithm or a comparable mechanism [3]-[6]. Some portions of such network, however, may still experience congestion if the aggregated traffic loads vary. On the other hand the routing method employed introduces some operational complexity since it often needs sophisticated and complex mechanisms with signaling protocols involved. Significant research in OBS networks focuses on alternative deflection routing [7], [8]. In such a routing scheme the burst is allowed to be deflected dynamically to an alternative routing path in a node if it contends with another burst on the primary routing path. Indeed, the deflection based routing can partially enhance overall network efficiency as well as performance under low traffic loads but may escalate the frequency of data burst losses under moderate to high traffic intensities [8].

Another approach to the routing problem makes use of the optimization theory and a few works can be found in this area [9]. In OBS network a burst loss probability is the main primary QoS metric of interest which adequately represents the congestion state of entire network. An approximated form of the end-to-end burst loss probability [9-11], has a non-linear curving which may make formulating the optimization problem quite complex.

1.2 QoS Challenges

The complexities and challenges in routing affect QoS levels and consistency thereof that can be guaranteed to the diverse traffic classes. The current lack or inadequacy of buffering facilities possess challenges in the running of OBS networks, especially, in situations where consistent QoS guarantees are desirable. Application such as real-time voice or interactive video have stringent QoS requirements and thus require additional QoS differentiation mechanisms in order to be protected from non-real time data traffic especially when the network resources are fairly constrained.

In this context mechanisms that offer minimally low burst blocking probabilities, delays as well as delay variations will be key in the successful roll-out and operation of OBS networks. Various QoS differentiation and implementation approaches have been investigated quite exhaustively in several literatures e.g. [12], [13]. Primarily, two QoS differentiation models are defined: absolute versus relative differentiation [12-14]. With the relative QoS differentiation model traffic is categorised and segregated according to defined threshold classes. Each class is not explicitly defined quantitatively in absolute terms, but instead, the QoS of each is defined relatively. Absolute QoS model generally aim at the provisioning of worst-case guarantees on the delay, loss, as well as required bandwidth to services. This type of stringent guarantee is regarded essential for the classes of real-time as well as loss sensitive services examples of which include mission-critical services.

Overall, QoS differentiation can be provisioned with respect to forwarding performance) or rather according to service availability. QoS differentiation ensures predefined QoS guarantees during a normal, fault-free network operation while the QoS Absolute model provisions and ensures QoS-enhanced protection mechanisms so as to afford and ensure resilience. QoS involves both the specifying of specific QoS classes, as well as incorporating dedicated measures to such classes [12]. Furthermore, individual classes are specified by pre-fixing upper limit bounds parameters such as burst loss probability, jitter as well as end-to-end latency. The delays arise mostly due to the introduced offset time, optical fibre delay lines (FDL) buffering, the propagation delay in fibre links, and ingress node processing. Properly setting up the maximum distance allowed for the routing algorithm will help limit the first two factors. Furthermore, the delays incurred in the ingress node can be controlled by utilising timer-based burst assembly strategies. Further, the optical buffering, which generally has limited application in an OBS network, introduces relative delays. Since there are numerous factors that influence the round trip delays the problem of congestion and contention is more complicated and needs more focus. Whichever approach is taken; the ultimate goal must be to ensure an end to end QoS guarantees and consistency commensurate with modern user's expectations. The key to such an achievement is implemented as a congestion and contention minimization strategy to eradicate delay and burst loss probability hence improving the QoS in OBS networks.

1.3 Congestion and Contention in OBS networks

The need to ensure maximum network capacity, bandwidth allocation/traffic prioritization, and active fault measures; all of which serve as a few of many resources afforded to carriers in order to ensure service assurance and delivery was the main impetus towards the development and implementation of the NGN concept which promises to replace existing legacy networks as well as provide a converged service platform between fixed and mobile communication infrastructures. However as discussed in the previous section NGN or OBS networks are an integration of features of OPS and OCS networks, the transmission approach in this network is that data packets are deployed from the edge nodes without any acknowledgment of arrival or resources reservation this makes the network prone to data loss. Amongst the other factors affecting OBS both congestion and contention have proved to be the main contributors to increase in data losses.

1.3.1 Congestion

Congestion can be defined as an excess in aggregate demand of resources utilization as compared to the available capacity of the resources [2]. In OBS networks congestion is mainly experienced when the network is exposed to very high speed burst traffic. The current approaches and studies in congestion still need more improvement and this has been the main source of motivation towards this work. We need a practical control method to improve on the overheads on OBS networks switching capabilities with satisfying QoS constraints while doing away with congestion.

1.3.2 Contention

Contention is a process or condition that occurs at the switching level in a network node when more than one packets attempt to depart the switch from the same output port, utilizing on the same wavelength at the same time. In optical packet switching all data packets are in electrical form. In the event that this condition occurs, contention is resolved through the use of a random memory access in which the store and forward approach is employed. High priority packets are given preferential treatment by storing the low priority ones in RAM while the contending high priority packet is granted access to use the output port. The absence of buffering in OBS networks possess challenges with regards to contention handling. All data packets are all-optical including an optical switching layer that offers a finer service delivery that closes up on the gap between electronic and optical conversion. Currently a few approaches have been suggested for resolving contention in such networks. These include the use of fiber delay lines (FDL) and wavelength convertors (WCs). These will be discussed in detail later.

In small scale networks contention and congestion are effectively resolved through basic schemes using the three main approaches, namely, space deflection, optical buffering and wavelength conversion. In this work these approaches may not prove to be useful since the target network is a large-scale backbone network that serve vast geographical areas. According to [1], telecommunications operators utilizing unlike vendor technologies and networks have always experienced difficulties in applying their service and business requirements to the network and element management systems of the different network domains due to congestion and contention. It is thus suggested that better cost effective approaches be implemented in order to address the congestion and contention ultimately reducing delays, delay jitter as well as data losses in NGN.

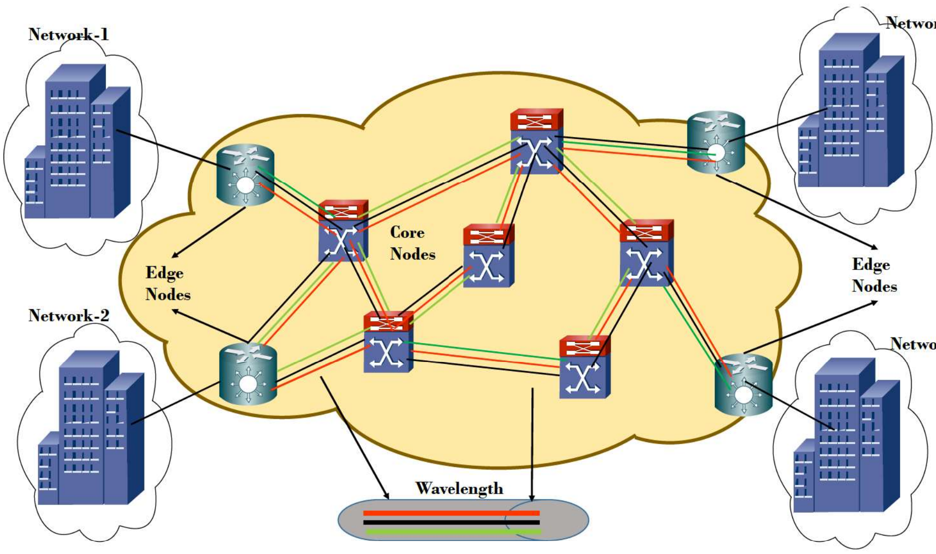

Figure 1-2: An Example Backbone Network.

Figure 1-2, illustrates a typical backbone network. The general architecture allows for an egress node at each regional client node. The data within the same edge node are aggregated to form a data bursts. Each burst requires a CP to be generated to lead and configure the resources for the burst that will follow at an offset time.

1.4 Resources Scheduling

Upon receiving the control packet, the core nodes reserve the requested resources based on the information contained in the CHP. A wavelength is quickly reserved and the switch fabric just in time to allow the burst to fly past it, a process also termed "fly-by". The resources reservation release can be either explicit of estimated. With explicit release, the source dispatches an explicit control packet to signal the conclusion of a burst transmission, whereas with estimate release, an OBS node is configured with details that state exactly the burst size and therefore can accurately calculate when the release of occupied resources can take place. During these operations, there is a greater chance of an increase in burst losses due to congestion and contention. This is mainly contributed to by the unavailability of resources while they are reserved for an unknown burst to be transmitted.

Overall proper routing and switching strategy coupled with effective network control strategies would help overcome and solve the congestion and contention problem of affording a consistent and guaranteed QoS in present day OBS networks. Routing control and management will enable network operators to efficiently and smoothly carry user data with desired QoS constraints. A further dimension in this regard is focusing on burst buffering through the application of fibre delay lines and other buffering based strategies. These will be discussed in subsequent chapters. In short, the future optical network switching paradigm and control plan is expected to be more versatile, scalable, automatic, and resource and energy efficient. Therefore, this specifically calls for further research on specific areas such as wavelength management, routing, control and management supporting multi-layer and multi-domain optical transport networks. Performance evaluation of control and management techniques, burst assembling algorithms, network reconfiguration as well as dynamic service provisioning have yielded promising results.

We conclude this section by noting that contention and congestion in OBS switched networks, is a major challenging problem primarily because of the limited buffering provisioning in the interior nodes. The buffer-less (or limited buffering) nature of OBS's score network switches readily attracts "burst contentions" and "burst congestion". As elaborated before, contention occurs due to multiple bursts contending for the output ports simultaneously, in which case only one burst will be successfully switched through whereas the rest will be served according to available contention resolution mechanisms. Congestion is when the available resources are unavailable for use by other services and it's normally triggered by traffic over load or switch equipment malfunctioning. This will in low utilization of some sections of the network as well as high burst drop rates, thus once again adding to QoS degradation and consequently its inconsistency.

Availing overall up-to-date information about the network bandwidth availability or constraints within an acceptable time scale may be a promising solution. E.g. in [15], [16] a distributed versus centralised framework approach is proposed. With the distributed method each network component has to have its own transport control entity. This includes a dedicated signalling engine for facilitating real time communication with other associated network components. The centralised approach uses a common signalling engine as well as a database. Consequently, the signalling associated data traffic within the network is reduces significantly. The actual network utility data formation is not distributed. For resource modelling of the overall network two approaches are possible. These are the complete comprehensive network model approach [15], and the proxy approach.

1.5 Overall Research Objectives

We will focus on analysing mechanisms and approaches geared towards alleviating congestion and contention in OBS networks, which by themselves are expected to form the core backbone platform for NGNs. As such we will specifically investigate new paradigms for routing strategies considering the fact that resources allocated to a user or service can be squatted temporally by others to support emergency or higher priority traffic. The strategies for dynamic resource allocation or switching under such scenarios are analysed, and applied accordingly to existing infrastructures in NG all-optical – networks (AONs), with the ultimate goal of ensuring a consistent and guaranteed end-to-end QoS.

Summarily the main objectives of this work are as follows: -

- To review OBS switching paradigm, partially in motivation it as the ultimate backbone network switching for NGNs.
- To survey burstification schemes taking into account the diverse QoS requirements of various applications.
- To study existing contention and congestion resolution approaches, as well as various signaling scenarios.
- To explore routing and wavelength assignment/allocation approaches in OBS networks.
- To propose and evaluate a congestion and contention minimization scheme that aims improving QoS of real time applications.

1.6 Thesis Organization

Chapter 2 gives an overview of optical switching paradigms. These are Optical Label Switching (OLS), Optical Circuit Switching (OCS), Optical Packet Switching (OPS) and Optical Burst Switching (OBS). This chapter clearly depicts the main factors that have contributed in considering OBS switching paradigm as the most preferred candidate for NGN. The OBS switching paradigm combines the best features of both OCS an OPS.

In chapter 3, we further elaborate on detailed architectural concepts involved in OBS networks. In this chapter our focus is on the functions and operations encapsulated in the OBS edge and core nodes and the overall performance structure of the OBS backbone network.

Chapter 4 surveys candidate traffic models for OBS networks. These models are then used in OBS traffic prediction and performance measurements. Generally, every channel in an OBS network when configured to transmit a large data burst packet it is made unavailable to other channel requests for a certain period until the data burst is assumed to have been completely transmitted hence in this chapter we also overview the time frames in which this blocking period(void) in OBS can be managed and how it also aids in contention and congestion build-up. A few traffic prediction models and concepts are reviewed in view of implementing them later on. In particular a selected approach will be used in our simulation scenarios in chapter 5.

Chapter 5 focuses on the main causes of congestion and contention in OBS networks. Further, we propose and describe a congestion and contention minimization strategy followed by its analysis. The analysis includes the modelling aspects as well as simulation.

Conclusions are drawn in chapter 6.

Chapter 2. OBS Networks Overview

Recent developments in the areas of IP, mobile technologies, and data center networks, cloud computing and social networks have resulted in the growth of a wide range of network applications as well as a surge in bandwidth demands. The data rate of these applications also varies from a few megabits per second ($Mbps$) to Gigabits per second ($Gbps$) ranges. To support this growth in data traffic, one foremost solution is to utilize the advancements in optical networks. With multiplexing technologies WDM, bandwidth in the ($Tbps$) ranges can be exploited from the optical bearer in an energy efficient manner thus at the same time promoting energy efficient networking. The OBS switching paradigm is currently the ideal choice for addressing the switching bottleneck, hence in this chapter, we provide an overview of a few optical switching paradigms and mitigate on why OBS currently emerges as the candidate ultimate switching solution for future generation networks.

2.1 Evolution of Optical Transmission Networks

While the idea of all-optical networking had its origins in the research community quite long ago, its practical realization was always is doubt. The original aim of the AON was to maintain the data wholly in optical domain end-to-end so as to eliminate E/O and OE, associated bottlenecks, and to allow transparency. Currently, backbone network traffic has surged. This has led to poor QoS in the transmission of data and thereof leading to a demand for a more sophisticated network approach that can sustain the bandwidth capacities at both transmission and switch level of data. Optical transmission is seen as the best for provisioning the required bandwidths. Optical networks have revealed, proved and continue to win multitudes of network engineers awareness of the high potential impact they have in providing the required bandwidths in future generation backbone networks. A few optical switching paradigms which include, OCS, OPS, and OBS are reviewed in the next sections.

2.1.1 DWDM Based Optical Packet Switching (OPS)

In OPS, individual data packets are transmitted individually end-to-end in the network directly in optical domain [10]. Its primary objective to allow for efficient usage of available network resources, namely bandwidth through statistical multiplexing since switching is performed at the packet level. Such a network is depicted in Figure 2-1. A CHP is appended and thus transmitted with the data payload. At each switching node (packet switch), the payload is temporarily stored in the fibre delay lines (FDLs) whilst the header is electronically processed. This processed CHP information will be used to configure the switch (node) accordingly. The FDLs thus play a crucial role, to allow/facilitate the electronic processing, as well as fabric reconfiguring that follows afterwards. Before the packet is sent out to the next node, its header is updated. OPS is regarded as a good switching paradigm when compared to other optical switching methods.

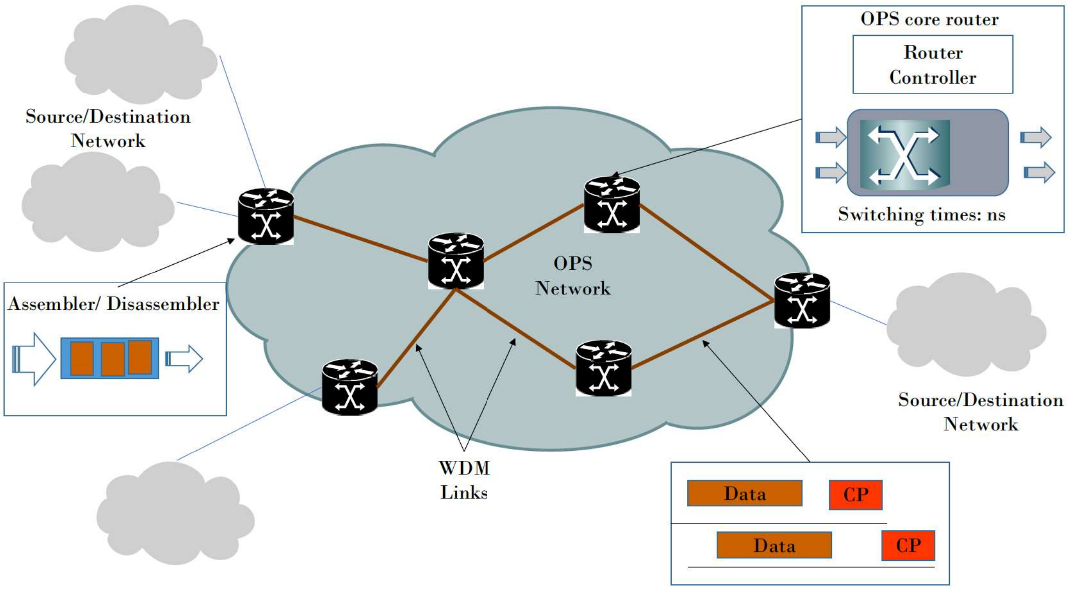

Figure 2-1: Example OPS Based Network

Because the IP based Network is designed for as a packet (datagram) switched rather than circuit switched, this approach makes OPS preferential in comparison with OCS. However its short comings are quite pronounced. Prominently is the fact that practical implementation of OPS is still not feasible.

Practically OPS would be very difficult to realise at this time as its implementation would require among other things, high speed (in orders of nanoseconds) switch fabrics, optical buffers, wavelength converters, header recognisers and other related components for the entire configuration of control and burst packets transmission.

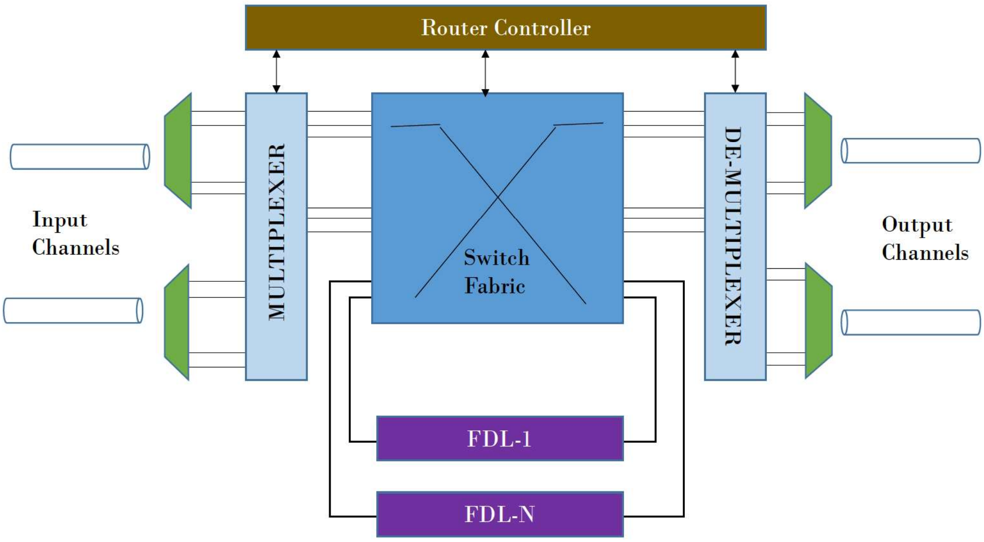

Figure 2-2: Router architecture of an OPS Node

Figure 2-2 shows the architectural operations/processes on each router in an OBS network. In terms of processing this includes multiplexing and de-multiplexing, O/E and E/O conversion and vice-versa, packet header extraction, packet header re-insertion, and switch reconfiguring at each network node.

2.1.2 DWDM Based Optical Circuit Switching (OCS)

In OCS all data is carried dedicatedly from one input wavelength to another at the switch's output port, Thus there is no need for O/E and E/O conversions required with this approach as is the case with OPS; hence there is no demand for overhead processing. A light path has to be established end to end i.e. from source to its final destination, with added continuity constraint [5].

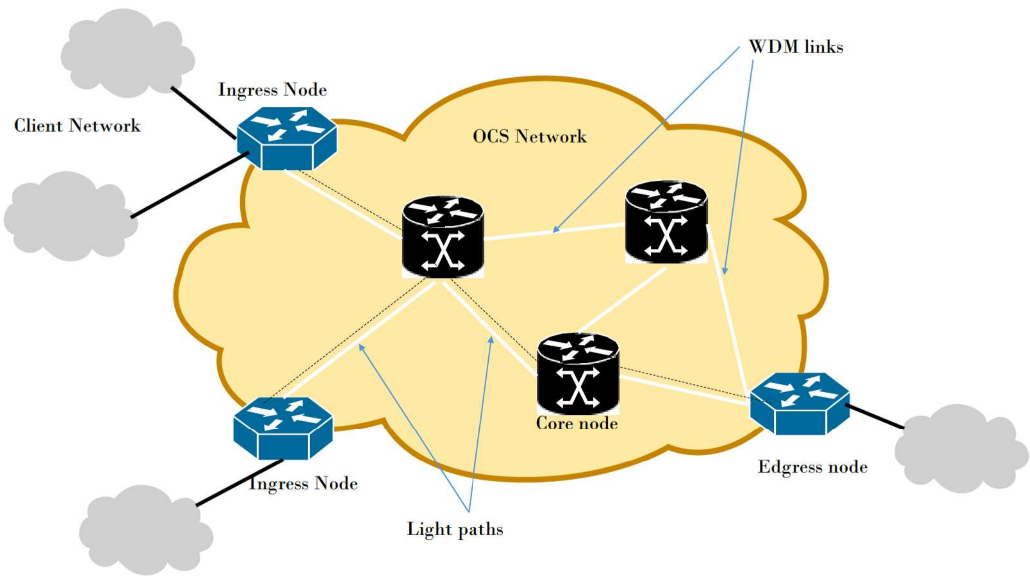

Figure 2-3: Illustration of an OCS Network

A switching node requests a light path to be established to the destination node via signalling all associated intermediate switching nodes (i.e. Along the path to the destination nodes). Once the setup phase is complete the data exchange follows in which the lightpath is exclusive to the two parties (sender and receiver switch nodes). An OCS network is illustrated in Figure 2-3.

The benefit of this approach is that of no requirements for any wavelength conversions as well as optical buffering. Disadvantages include the "Routing and Wavelength Assignment problem" [7]. Its solutions are dealt with in [8] and [9]. Herein the authors highlight the high blocking probabilities experienced. They suggest the use of wavelength converters (WCs). A wavelength converter transforms an incoming wavelength to a different value so as to avoid blocking. Its use however, has the repercussions degrading the signal [9].

2.1.3 Optical Label Switching (OLS)

Reference [11], [12] propose Optical Label Switching (OLS). With OLS, when a packet is received at an ingress node, the routing information is extracted from the header and placed in a special tag called a label. From this point on, the payload and label are served separately. As both traverse the network, the label is processed in electronic domain at each node for routing decision purposes. It is worth noting that the payload will remain in optical form through the entire network. Below is a depiction of data payload and optical header/ label coupling /de-coupling and further streaming.

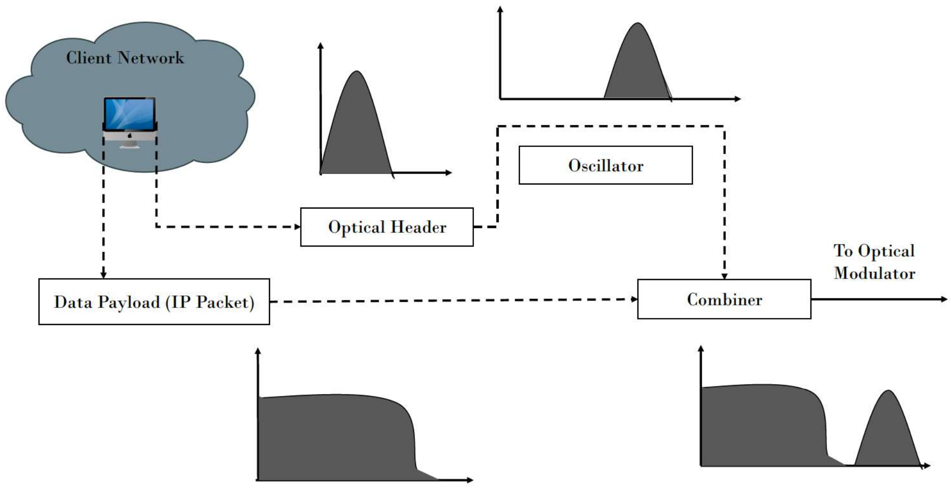

Figure 2-4: OLS header/ data payload coupling/decoupling

When data packets are received at an ingress node, they are assembled into optical payloads. The assembling of the data packets is according to destination address, and associated desired QoS parameters. Each payload requires a separate label.

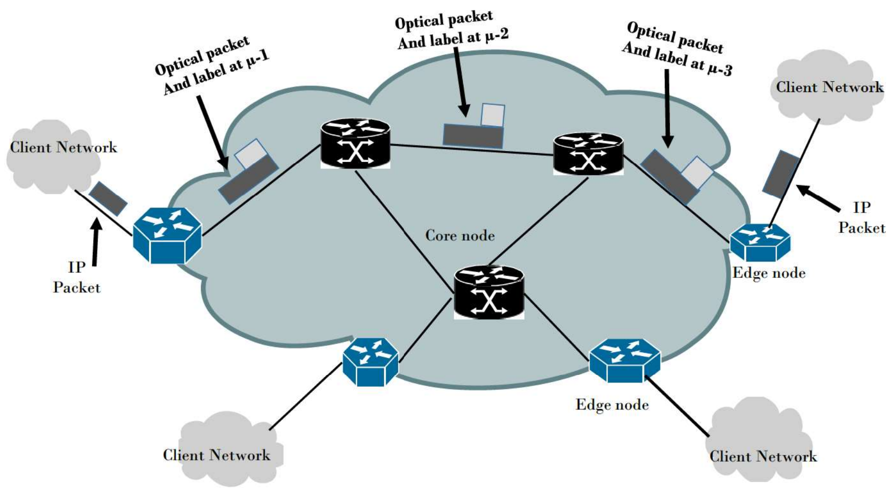

Figure 2-5: Optical Label Switching Network

Both the label and payload are then together transmitted. The two crucial components of an OLS network are core and edge nodes. Edge nodes provide interfacing services between the customer and OLS transport networks. Furthermore these ultimately detach the label from the optical payload and decouples the payload to revert to the original data packet format. Ultimately the packets are delivered to their respective destinations. Figure 2-5 shows an OLS network architecture.

2.1.4 Optical Burst Switching (OBS)

In OBS, relatively a few CCs (e.g. one per fibre) go through O/E/O conversions [5]. Since the data is served wholly in optical domain and at burst level, data transparency as well multiplexing efficiencies are achieved concurrently. As cited previously with OBS the CPHs are separated from data payloads right from the ingress nodes. Traffic (packets) destined for a particular egress switch are assembled into one huge data packet called a "burst". The CPH is then sent prior to its burst at a carefully chosen offset time. This is to afford the CPH sufficient time to pre-configure the next switch so as to afford fly-by routing of the burst upon its arrival.

The CPH and the burst are sent on separate dedicated channels. OBS takes advantage of both the ultra-bandwidth availability of a fibre for switching/transmission and the advanced data processing capability of modern VLSI electronics. It can achieve a sizeable cost reduction and leverage

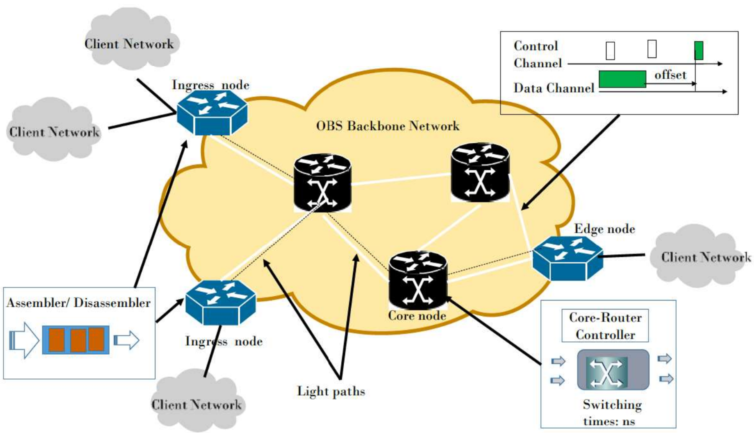

Figure 2-6: Optical Burst Switching network

At any given OBS node, no alignment of bursts is necessary unless if the node's switching network is slot operated. Furthermore, FDLs and WCs which are optional can aid in lowering burst losses [13]. At present, it is a task to implement an OBS switching network with multitudes of ports operating gigabit per second speeds. Advances in research in this regard is ever improving the required speeds [13-15].

2.2 OCS, OBS and OPS Comparisons

OPS and OBS have similarities. However their primary difference is that in OBS, data packets are assembled into bursts while with OPS data packets transmitted individually. On one hand, OBS control information (CHP) is signalled on a dedicated channel while in OPS it is signalled in-band, i.e. the control information is appended to the packet itself. Furthermore OPS will require O/E/O conversion and vice-versa of all data transmitted whereas with OBS only CHP is processed electronically by it undergoing O/E/O conversion. Other characteristics that are deemed important when comparing optical switching paradigms include; hardware requirements, flexibility, performance, processing complexity, and QoS. These are discussed next.

2.2.1 Processing Complexity

OBS and OPS both require specific algorithms for contending with contention resolution situations. They consume compound algorithms for QoS provisioning so as to be able to exploit benefits of statistical multiplexing. These contribute to the expansion of OBS/OPS processing complexity of control algorithms. Unlike OCS where light-path is calculated and established once before transmission, in OBS/OPS calculation of the path is complicated simply because it is done at each intermediate edge node. OBS switch control has to maintain a precise timing of a burst packet payload arrival relative to its header packet arrival. The header packet should setup the path before the data burst arrives at each edge node, hence the offset times in OBS becomes more complex to maintain.

2.2.2 Performance

The performance in this case is discussed by considering network throughput, utilization, transmission delay and burst packet loss probability. [18] [19] [20] OCS performs incompetently compared to OBS or OPS in terms of transmission of statistically multiplexed traffic hence network utilization is high in OBS/OPS compared to OCS because statistical multiplexing permits a better exploitation of network resources. As far as throughput is concerned, OBS/OPS have a problem of contention which gives a negative result simply because of the unavailability of optical random access memories (RAM). OPS engages FDLs buffering and WC to alleviate the challenges associated with packet loss probability. Another concern in this switching paradigm is the Transmission delays. This is mainly due to continuous processing of data packets at intermediate nodes and the connection setup. Other delays are mainly attributed by the switching technologies used. In OPS delay is very small or rather negligible. Alternatively OBS has a significant delay due to a significant time spent during the process of burst assembling. In OCS the delay is due to the time needed for the establishment of a lightpath.

2.2.3 Hardware Requirements

Due to the dynamic nature in switching time requirements, each optical switching paradigm requires different and appropriate switching technology. E.g. Burst switching with end-to-end setup requires speeds in the magnitude of milliseconds, in that case micro-electro-mechanical systems (MEMS) would be the best available option.

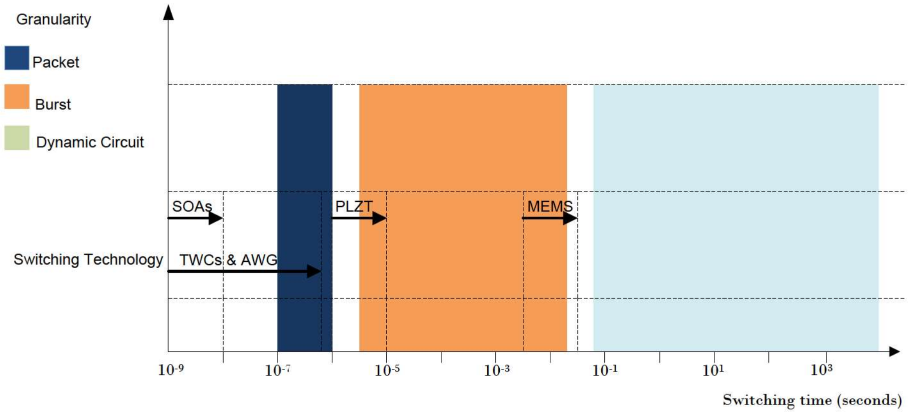

Figure 2-1: Switching granularity and required technology

The best choice for optical burst switching with microseconds switching times is Lead lanthamun Zirconate Titanate (PLZT) technology. For packet switching with Nano-seconds speeds, Arrayed Waveguide Grating (AWG) and tunable wavelength converter (TWC) technologies or semiconductor optical amplifier (SOA) technology [16] [17] are used. Figure 2-6 shows switching times for circuit, burst and packet granularity and the required technologies.

2.2.4 Flexibility

Flexibility here is regarded as the ability of the switching paradigms to adapt different bit-rates and data formats to transmit them in one network. It is one of the important factors to consider when comparing any switching paradigms in computer networks. It is mainly achieved by having proper functions in edge nodes. In OBS adaptation is a difficult task. However, granularity in OBS and OPS can be flexible compared to in OCS. In OCS paradigm, this is achieved by the use of E/O/E conversion. One of the advantages of OBS/OPS over OCS paradigms is that they are more scalable.

2.2.5 Switching Comparisons

OCS has no demand for any QoS consideration. This is mainly due to the fact that its resources are assigned at ingress nodes prior to transmission process. OBS/OPS demand for QoS implementation because each burst/packet has certain QoS attributes that contribute to its entire transmission process. That means each burst/packet needs to be processed depending on what QoS attributes are encoded or embedded in it. QoS will be discussed in detail in section 3.6.

A comparison of the three switching paradigms is summarized in table 2.1

Table 2-1: Comparison between OCS, OPS and OBS[17].

Criteria	OCS	OPS	OBS
Processing Complexity	Low	High	Medium
Setup Latency	High	Low	Low
Bandwidth Utilization	Low	High	High
Granularity	Coarse	Fine	Moderate
Traffic Adaptability	Low	High	High
Switching Speed	Slow	Fast	Moderate

2.3 Basic Architecture

OBS has become a practical solution towards narrowing switching versus transmission speed gaps in future generation backbone networks. At transmission level, data packets sourced from edge nodes are aggregated and assembled into optical burst units generally referred to as bursts. A burst control packet is transmitted for each assembled burst in a dedicated control channel and delivered with a small relative offset-time prior to the actual data burst's arrival. This offset timing allows for *electronic processing* of the control packet by a controller at the next and subsequent nodes thus creating an allowance for a wavelength reservation on its output link and switch matrix reconfiguring usually for the duration time of the incoming burst and eradicating the need for burst buffering. The burst will then shortly fly by and immediately afterwards the reserved wavelength can now be freed /released and made available for other connections. This effectively alleviates the need for optical buffering at intermediate nodes which otherwise would escalate network design and operational costs. Furthermore, such a temporary usage of wavelengths promotes higher resource utilization as well as better adaptation to highly variable input traffic in comparison to optical circuit-switched networks. OBS architectures with limited buffering capabilities would still be susceptible to congestion states and in many cases exasperate it [1]. The existence of a few highly congested links may seriously aggravate overall network throughput..

2.4 Challenges in OBS Networks

OBS networks have proved to be the most promising switching technique for future generation networks as they are expected to accommodate the ever surging bandwidth demands in backbone networks. This architecture suffers from certain aspects that can hinder the success of the OBS networks. Ultimately in this work, we will focus on addressing the minimization of congestion and contention which together can lead to degradation of network performance.

2.4.1 Blocking Probability

OBS data packets arriving at a given ingress node, and intended for a common destination are aggregated into a series of huge data bursts, each being switched and routed independently. In this way, the switching or configuration processes on the network are simplified since only one header associated with each burst is used to configure the desired channel for a burst packet. Therefore the CPH precedes the burst payload and reserve the required resources at each node's fabric and associated output link port along the route. The payload or burst packet then shortly follows the CPH without waiting for an ACK after a staggered offset time. At every node, if the requested bandwidth is available, the burst is transparently switched to the next intermediate node, otherwise, the burst is discarded A measure of the blocking rate is quantified by what is termed "the blocking probability".

2.4.2 Congestion

Resources utilization is one of the main key factors of the preferences associated with OBS networks. The main challenge is that in some or rather most cases the deployment of burst packet and their corresponding control packet will generally increase to a rate whereby the switching demand in these networks will or may not be able to accommodate deployed data packets. In other words, the network would be suffering from an increase in demand for resources utilizations as compared to the available resources. Generally, this will give rise to the queuing of data packets waiting for the release of the network resources. By resources, we referring to the switches, routers, wavelengths etc. Since we highlighted that burst packets are transmitted without any acknowledgement, basically congestion leads to increases in burst loss, jitter delay, reduced QoS.

2.4.3 Contention

The finer granularity offered by OBS networks is the main cause of the reduced O/E/O conversions. In this paradigm, the light path allocation will mainly be done at each output node in the backbone network. Contention occurs at the switching level or node whereby more than one data packets are contesting to exit the same output port, on the same light path. In the basic electrical conversion, this is mainly resolved by the use of electronic buffers but in the all-optical domain, there is no equivalent optical RAM.

2.5 Summary

The main advantages of OCS are that it affords a natural acceptable and reliable QoS. It's relatively a mature technology and as such its components and subsystems are commercially available. Its problem is that of low flexibility as well as low network utilization. OCS have very high wavelength consumption and requires that the node be of large sizes. These key aspects rule it out as a switching paradigm solution for NGNs. Both OPS and OBS offer better flexibility and network utilization. However, OBS is able to offer efficient network utilization using lower-speed switching elements as compared to OPS. The biggest point of difference between OPS and OBS is the use of E/O/E conversion concept used by OPS and the concept of burstification of packets used by OBS. The E/O/E conversion and the need for optical buffers render OPS difficult to implement. OBS's burstification concept also contributes delays. However, both have high resilience and control complexity. OBS is considered a viable solution for NGNs as it does not require any new technologies. OPS would have been also a better option; however, it is still awaiting technological breakthrough especially for compact and low-cost optical components. It has a longer-term deployment and OBS has a midterm deployment. In this document we present optical burst switching as the ultimate solution for Next Generation Networks.

Chapter 3. Burst Assembling, Signalling and Scheduling.

OBS networks comprise collections of edge and interior nodes. The edge (ingress) node assemble user packets into a burst which traverses the OBS interior all the way to the destination edge (egress) node. The ingress node primarily processes and schedules data packets into input buffers according to desired class and ultimate sink address. Further the packets are mapped into bursts. These in turn will be transited wholly in optical domain over the interior network, without any necessities for temporary buffering at intermediate nodes. In chapter two, we compared switching paradigms such as OCS, OPS and OBS. In this chapter OBS concepts such as burstification, scheduling and signalling are reviewed.

3.1 OBS Node Architecture Overview

As illustrated in Figure 3-1 an OBS network mainly consists of edge and core nodes. These are interconnected by DWDM links.

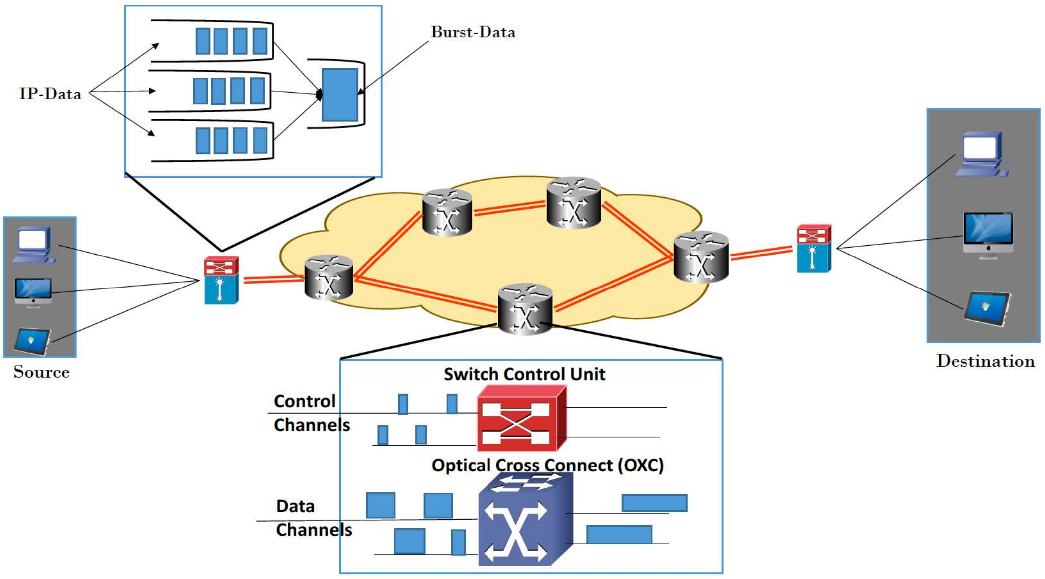

Figure 3-1: OBS router/switching architecture

In the previous chapter, we discussed the technical processes involved when IP data packets from access networks arrive at the ingress or edge nodes. Basically, IP data packets are aggregated and assembled into optical bursts data packets and at the same time a corresponding CPH for network resources configuration is generated for the burst. In this case there exist various crucial approaches for IP data packets aggregation and assembling. In this chapter we focus on these crucial aspects and the impact they impose on the overall performance of optical burst switching.

First data payloads destined for a common egress node are aggregated into a burst whose size is determined by the chosen assembly algorithm. Bursts destined for a common destination can also share a single header packet. The header packet contains key information for configuring and routing. Key information includes burst length, classification, destination as well as QoS attributes to be put into consideration. Figure 3-2 shows the generic format of the control packet or header packet in OBS networks.

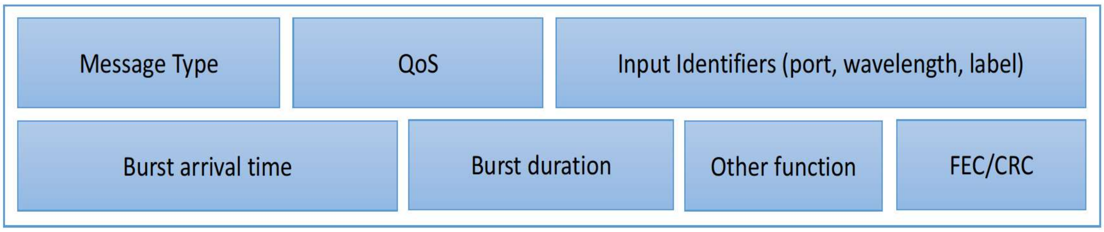

Figure 3-2: Typical Burst header Packet format

The CPH is transmitted ahead of its associated data burst packet with a given offset time. This is to allow for the configuration process of the switching fabric in the core nodes. The data burst and the CPHs are sent on separate dedicated channels. Data burst payload is transmitted all-optically while header packet undergoes the previously discussed O/E/O conversions. The CPH is processed by an electronic controller of the switching or node in the interior. The controller performs several functions including contention resolution, QoS provisioning, burst forwarding, resources reservation,, etc. These operations will be discussed in subsequent sections.

3.1.1 OBS Edge Node

OBS network consist of two main types of edge nodes, thus, the ingress and egress edge nodes. Each edge node consists of a set of operations that it executes based on its location over the network. The ingress edge node is obligated for aggregation of remote client data into one single burst packets. It then assembles several packets destined for a common egress node into a larger burst and also automatically generates a corresponding burst control packet. The ingress is also responsible for setting up an offset time between the CPH and the burst data payload. These functions will be discussed in the next sections. Figure 3-3 shows a typical OBS ingress node. The egress node disassembles the burst data payload into the original IP unit data packets before redirecting the data to the intended client networks. In actual sense the operations performed by the egress node is the direct opposite of the ingress node.

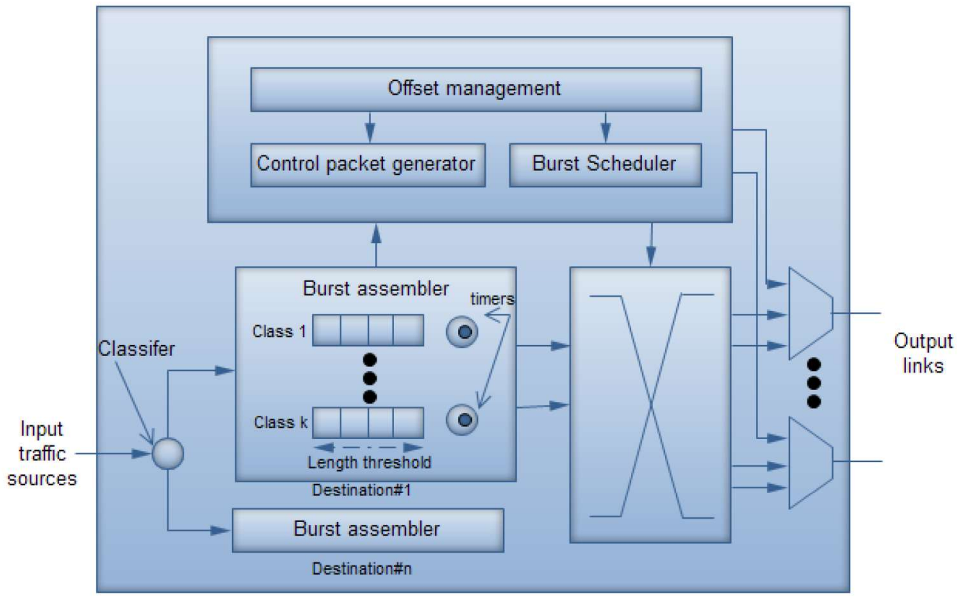

Figure 3-3: Typical OBS ingress node

3.1.2 OBS Intermediate Core Nodes

An OBS intermediate core node is located in the core of the OBS network. It consists of the following functional elements; an input interface, switch fabric, an electronic switch controller and an output interface. This is shown in Figure 3-64. The major role of the input interface is to perform the extraction of network configuration properties and data channels from the header packets. The data extracted from the control channel carry control information about the incoming super or burst packet. Only control information is converted form O/E. The component used to process header packets is the switch controller. Most importantly, the switch controller consists of a burst forwarding look-up table with all reservations for incoming data bursts or payloads. The actual resource reservation is discussed in detail in section 3.3. Resource reservation calls for contention resolution as well. The role of the output interface is to update control information, multiplexing of data and control channels.

In the following section, we will discuss burst assembly mechanisms/ algorithms used for the burstification process. Burstification is generally one of the fundamental functions performed by the ingress node.

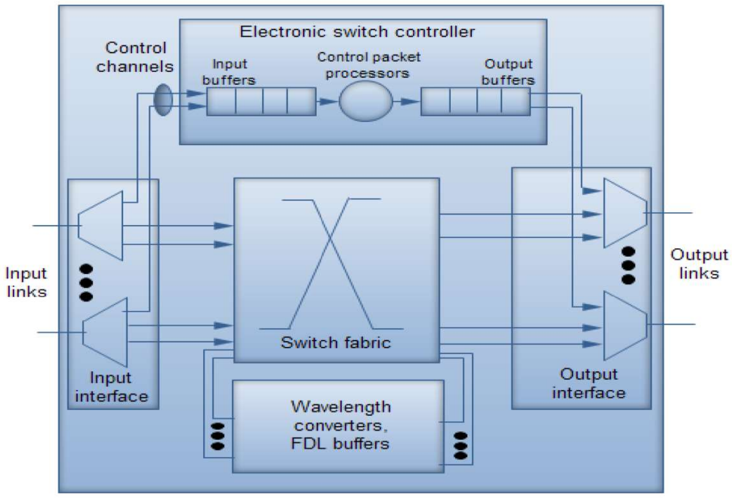

Figure 3-4: OBS intermediate node

3.2 Burst Assembling

Burst assembling is a process whereby incoming IP data packets from various client network sources are aggregated into a number of large data bursts at the edge node of an OBS network. As shown in Figure 3-6, the switching unit forwards the arriving packets to burst assembly units. The aggregation is performed using various criteria thus for example, packets which are bound for the same output lines are processed and aggregated into one burst assembly unit. The burst assembly criteria used influences the traffic characteristics in the network such as traffic burstiness and traffic self-similarity. The strategy implemented also determines the end to end performance of the network. There are basically two main objectives associated with the burstification mechanism thus;

- Reduce burstification process delay thus reduces the overall network data transmission delay.
- Increased the burst length – increasing the burst length will result in the reduction of the number of bursts produced and CPs, hence a reduction in the associated processing of burst packets and CP configuration overheads at the core nodes.

Figure 3-5 illustration on the performance of a "good" and a "bad" burst assembly mechanism/algorithm. These two objectives associated with burst assembling contradict with one another thus increasing the burst size will also increase burstification delays. It is, therefore, a challenging task for one to strike a balance between these two objectives when implementing an effective feature of burst assembly strategy. According to [21], "choosing the desired balance between the burstification delay and the burst size depends on the QoS requirements of the users, and the processing and buffering capabilities of the backbone nodes".

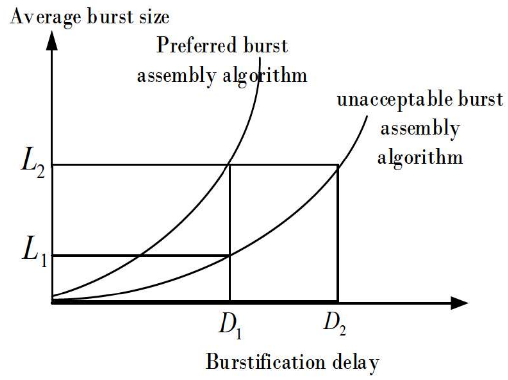

Figure 3-5: Burst assembly algorithm Performance

A number of significant burst assembly algorithms have been discussed in the literature. The traditional or rather basic assembly algorithms are mainly burst length-based (size threshold based) algorithms and time threshold-based algorithms [6]. Another algorithm which was proposed is the hybrid time/burst length algorithm which consists of mainly the properties of the individually stated algorithm. Other basic differences associated with these algorithms were also presented in the literature [21]. In recent years there have been quite a number of proposals of burst assembly algorithms which are based on traffic prediction/forecasting. In the next sub-section, an overview of these algorithms is given. The architecture of an OBS node at ingress is shown in Figure 3-6.

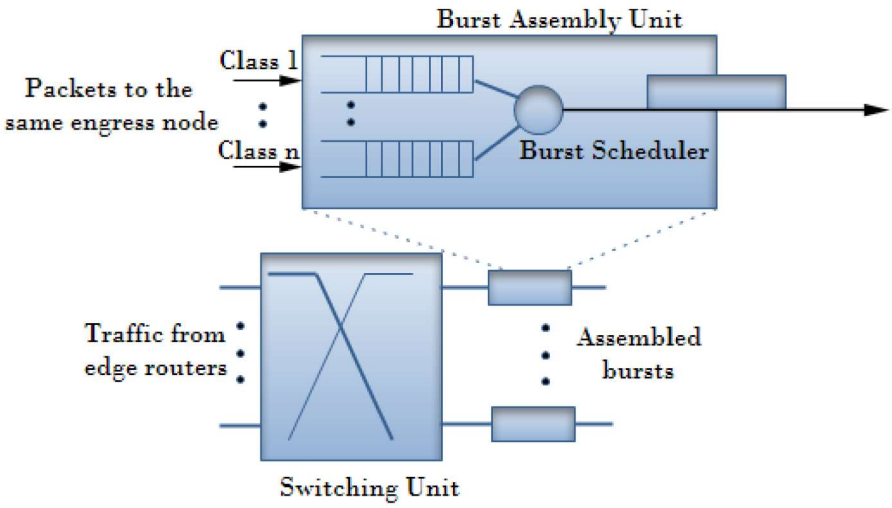

Figure 3-6: OBS Edge Router Architecture

3.2.1 Traditional Burst Assembling Algorithms

The following are typical traditional burst assembly algorithms used in current OBS backbone networks.

3.2.1.1 Time-Based Algorithms

In time-based algorithms, the fundamental feature is that a timer is set at the start of each burst assembly cycle. When the configured fixed time T elapses, all packets that would have arrived during that period would be assembled into a burst. This approach also consists of a minimum burst size. Therefore, all bursts generated should be of a size greater or equal to L_{min}. If the burst size is less than L_{min}, then it is padded to L_{min}. Careful consideration should be made when choosing the value of time T as large values may lead to unnecessary packet delays at the edge nodes. On the other hand another problem associated with a small value is, many small bursts will be generated in a short space of time and that would lead to an increase in control packets overhead at the core nodes [22]. The algorithm of time-based algorithm is as follows;

Step 0: While timer < T

Accept new arriving IP packet

Check timer

Go to Step 0.

Step 1: Check burst_size

If burst_size < L_{min} *then*

pad it to size L_{min}

Generate burst and send it output port

Reset timer

Go to Step 0

3.2.1.2 Burst Length Based Algorithms

Burst length based algorithm also known as size threshold algorithm was discussed in [6]. Contrary to the time-base algorithm, this algorithm uses a fixed burst size L to decide the generation of a burst packet. Once the length L configured on each edge node has been reached a burst is assembled and sent to the corresponding output port. Just like the time based algorithm this scheme also has its own short comings depending on traffic, it may also take a significant time for L to be reached and as a result an increase in delays at the edge node may be experienced on the OBS network. Again in the imminence of high traffic volumes, with a smaller configuration of length L, many bursts will be produced resulting in high control overhead at the core nodes.

3.2.1.3 Hybrid Algorithms

In order to deal with the overheads associated the previously discussed schemes in section 3.2.1 and 3.2.2, a hybrid algorithm was proposed in [6]. In a hybrid scheme a burst is assembled when either L or T is reached, whichever comes first? When T is reached first and when the burst size is not equal or more than the set L_{min}, then the burst size is padded to L_{min}. [22] The algorithm is implemented as follows;

3.2.2 Other Assembly Approaches

In an arduous search to further improve the performance of the burst assembly described above, several variants of the above burst assembly algorithms were proposed. In this section a brief description of some of them is given.

3.2.2.1 Adaptive-Threshold with Fixed Maximum Time Limitation Algorithm (ATH-FMTL)

The Adaptive-Threshold with Fixed Maximum Time Limitation (ATH-FMTL) was proposed with the view of providing flexibility to the actual network traffic. In the previously discussed algorithm which lacks in the traditional burst assembly algorithms in section 3.2.1. The main aim of the ATH-FMTL algorithm was to achieve a point of equilibrium between incoming IP data packets and the transmission of the bursts over the network. That means the rate of packet arrival at the ingress node corresponds to the rate at which bursts are formed and transmitted.

ATH-FMTL algorithm capitalises on the optimal burst length threshold and fixed maximum time limitation as features of burst creation. According to [23], "*The burst length thresholds are increased or decreased in case the burst queue size, at the time of burst generation, is larger than upper threshold or smaller than lower threshold, respectively.*" The function of ATH-FMTL is as follows: [23]

1. To classify the packets in appropriate burst, the criteria performed is based on the fact that every IP data packet has a delay tolerance that allows for flexibility during packet routing and on the assumption that no packet has a delay tolerance less than the amount of time it takes to route the packet through the OBS network, using the shortest route to its destination.
2. Also on packets arrive at the corresponding port and service classes these assembly queue becomes operative.
3. Each burst length is estimated at the end of prediction time (t_p) according to the previous burst length value and current arrival traffic.
4. Edge nodes determine the variable burst assembly duration (V_{bad}) by estimating burst size with current or previous load.
5. A CP is sent to OBS core network at time τ ;

   $\tau = t_a - t_0$ (t_a: assembly time; t_0 : offset-time)

6. It is assumed that wavelength conversion is available at every Label Switch Router (LSR) node in the core and a Just-Enough-Time reservation scheme is used.
7. The Latest Available Unused Channel with Void Filling (LAUCVF) scheduling scheme is used.

The pseudo code for ATH- FMTL algorithm is given below;

Begin
Packet arrives at node s;
d: destination of the new IP packet
If (Fixed assembly duration (Q_{sd}) is not running) then
Start VABDs timer for Q_{sd};
End if
Assemble the packet to corresponding queue Q_{sd};
Update the burst length information;
If (Assembled burst Length ≥ Expected burst length) then

Generate burst control packet for this burst;

Fill in and send out the control packet on a control channel;

Schedule the data burst to be sent out on a data channel after an offset time;

Stop the VABDs timer for Q_{sd};

The burst length margin is filled with void into the next data burst;

End if

End

Q_{sd} Denotes the assembly queue at the ingress node s for destination d and $VABD_s$ being the assembly duration used by the node.

In [23] it is concluded that ATH-FMTL does not only increase the switching efficiency at the OBS core nodes, but also smoothens the input packet traffic and reduce the data loss to a significant value. The main aspect identified as the one which can minimize loss is, being able to find an optimal threshold rage.

In [21], three burst assembly algorithms were proposed which are based on the estimation of the expected number of packets arrivals. At the end of each time frame, a decision needs to be taken as to whether the burst should be sent out and a new burst assembly started or the edge node should wait for more packets from the next frame. For such a decision to be taken, a linear prediction filter to estimate the expected number $\hat{N}(n+1)$ of packets arrivals in the next frame $(n+1)$ is used, to monitor if any specified criteria is fulfilled. The three algorithms based on this principle were proposed in [21] and are discussed next.

3.2.2.2 Fixed Additional Packets Threshold Algorithm (FAPTA) (N_{min} algorithm)

The FAPTA algorithm defines a lower bound $N_{\min}$. This is the bound on the number of packet arrivals above which a decision to wait for more packets before sending a burst packet is assembled is made. After frame n, the comparison with N_{min} is made to assess if it is smaller than $N_{\min}$, then the burst is assembled and sent with immediate effect or else it waits for another frame to be completed. In this scheme the burst is sent out at the end of the n^{th} frame if and only if;

$$\hat{N}(n+1) < N_{\min} \qquad (3.1)$$

3.2.2.3 Proportional Additional Packets Threshold Algorithm (PAPTA) (α L algorithm)

Unlike the fixed additional packets threshold algorithm this scheme makes use of a fraction of the current burst length $L(n)$ as the threshold value. The burst is assembled at the end of the n^{th} time frame if and only if

$$\hat{N}(n+1) < \alpha.\mathrm{L(n)} \qquad (3.2)$$

Where $L(n)$ is the burst length, $\hat{N}(n+1)$ is the schemes estimate of the number of packet arrivals expected during the frame $(n+1)$ and α is the multiplicative parameter.

3.2.2.4 Average Delay Threshold Algorithm (ADTA) (T_A algorithm)

The ADTA Algorithm was also proposed in [21]. The main goal behind this algorithm was to explicitly improve the Average-Delay-based algorithm which was proposed in [24]. With this scheme the running average delay is calculated and the burst is assembled as soon as the average delay of the packets reaches the threshold T_{AVE}. A couple of draw backs are experienced in this algorithm [24] these are;

- The process of calculating the running average introduces network overheads.
- Bursts transmission may not be sent out with an optimal time.

In a quest to address these drawbacks, the T_A algorithm makes use of traffic prediction schemes. The burstification delay is estimated at the end of each frame. The burst is sent out when the estimate exceeds the threshold value T_A. The average packet delay, $D(n)$ of the packets in the burst assembly queue at the end of the frame is defined as;

$$D(n) = \frac{\sum_{i=1}^{L(n)} T_i(n)}{L(n)} \tag{3.3}$$

Where $L(n)$ the burst size at the end of frame n, $T_i(n) = n.\tau - t_i$ is the delay of the i^{th} packet. $D(n)$ is computed using recursion as follows;

$$D(n) = \frac{L(n-1).D(n-1) + L(n-1).\tau + \sum_{i=1}^{N(n)} T_i(n)}{L(n-1) + N(n)} \tag{3.4}$$

Where $N(n)$ is the number of packet arrivals during n frame. The estimated Average Packet Delay $\hat{D}(n+1)$ at the end of frame $n+1$ is given by;

$$\hat{D}(n+1) = \frac{L(n).D(n) + \tau.L(n) + \hat{N}(n+1).\frac{\tau}{2}}{L(n) + \hat{N}(n+1)} \tag{3.5}$$

A burst is assembled at the end of the n^{th} frame if and only if

$$\hat{D}(n+1) > T_A \tag{3.6}$$

Where T_A is the predefined threshold value.

3.2.2.5 Average Delay to Burst Size Ratio Improvement Prediction Algorithm (L_{MIN} algorithm)

The L_{MIN} algorithm makes uses traffic prediction to compute an estimate $D\hat{B}R(n+1)$ of Delay to Burst Size Ratio (DBR) at the end of frame $(n+1)$. The burst is sent if and only if the estimate is worse than the current $DBR(n)$. The $DBR(n)$ at the end of frame n is given by;

$$DBR(n) = \frac{D(n)}{L(n)} = \frac{\sum_{i=1}^{L(n)} T_i(n)}{L^2(n)} \tag{3.7}$$

or it can be calculated recursively as;

$$DBR(n) = \frac{L(n-1).D(n-1) + L(n-1).\tau + \sum_{i=1}^{N(n)} T_i(n)}{(L(n-1) + N(n))^2} \tag{3.8}$$

$D\hat{B}R(n+1)$ at the end of frame $(n+1)$ is found as follows;

$$D\hat{B}R(n+1) = \frac{L(n).D(n) + L(n).\tau + \hat{N}(n+1).\frac{\tau}{2}}{(L(n) + \hat{N}(n+1))^2} \tag{3.9}$$

The burst is assembled and sent out at the end of frame n if and only if;

$D\hat{B}R(n+1) < DBR(n)L(n) > L_{MIN}$

L_{MIN} Threshold is used as a lower bound on the length of the bursts.

3.3 Burst Reservation Protocols

A number of burst reservation protocols in OBS have been explored by researchers in a quest to determine effective approaches to improve OBS networks. In this section, we discuss the main parent approaches to this technique thus one-way and two-way reservation protocols.

3.3.1 One-Way Reservation Protocols

In this approach, when a request for resources reservation is sent there is no acknowledgement sent back to the sender by the recipient nodes. A good example of a one-way reservation protocol in OBS network is the Just-Enough-Time (JET) which was proposed in [25]. JET, the control packet is transmitted first ahead of the corresponding burst with a set offset time t. This approach cuts down on the overheads of sending an acknowledgement on the resources reservation. The approach leads to low utilization of the edge router's access link. Another example of the one-way reservation protocol is the Just-In-Time (JIT) [26], [27]. JIT uses the same principle as JET with a slight difference. It only uses an acknowledgement from only the first cross-connect. Since both JET and JIT do not use acknowledgements, some bursts might be lost if they are sent and their transmission requirements are not met. Therefore, one-way reservation protocols have a high blocking probability. Illustrations of one-way reservation protocol and JIT are shown in Figure 3-7 and 3-8 respectively.

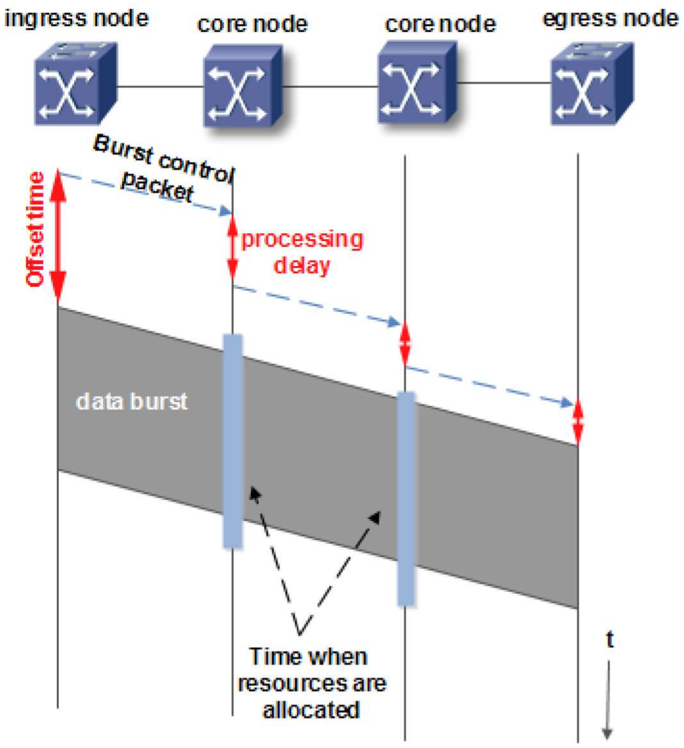

Figure 3-7: Typical One-Way Reservation Protocol

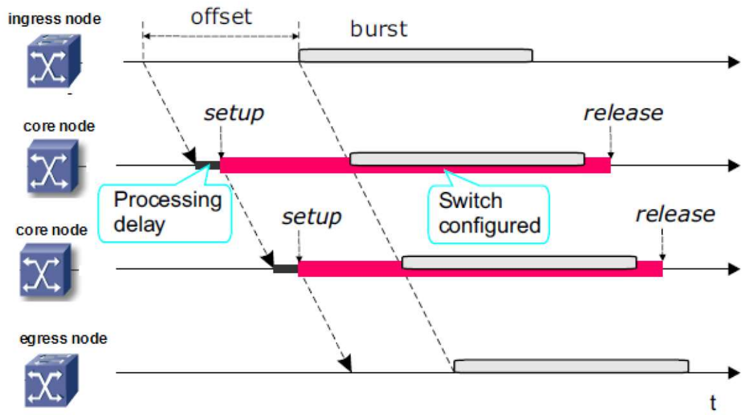

Figure 3-8: Typical Just-In-Time Reservation Protocol

3.3.2 Two-Way Reservation protocols

Unlike one-way reservation, in this approach an acknowledgement is sent back to the requesting network resources. The main important aspect of this scheme is that when a request has not succeeded the burst will remain in memory until the request is confirmed its succees. An example of a two-way reservation protocol is wavelength-routed optical burst switching (WR-OBS) [28]. One of the main features of optical burst switching is burst scheduling. Section 3.4 discusses burst scheduling.

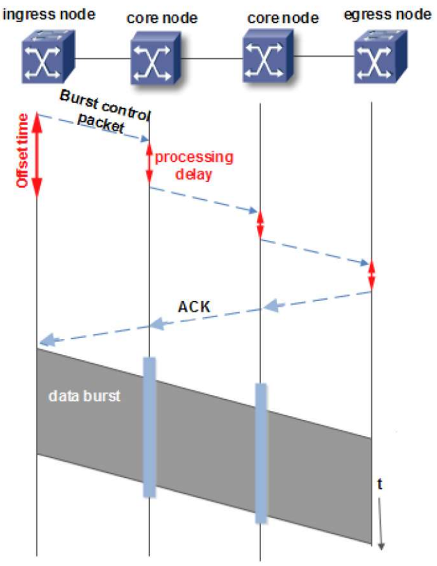

Figure 3-9: A typical Two-Way Reservation Protocol

3.4 Burst scheduling algorithms

Scheduling algorithms are mechanisms used to decide on which wavelength to assign an outgoing burst. If there is no wavelength available, the burst scheduler uses other tools at its disposal to delay the burst transmission, for example, fibre delay lines (FDLs). There are two main categories of burst scheduling algorithms. There are those which are based on JET reservation protocol and those that are based on Horizon. JET reservation protocol was discussed in the previous section thus 3.3.1. In [29] *"The scheduling Horizon is defined as the latest time at which the wavelength is currently scheduled to be used"*. The horizon scheduler selects the wavelength whose scheduling horizons are smaller than the burst's arrival time. The scheduler then updates the scheduling horizon to be equal to the time when the burst is due to be completed [29]. Horizon scheduling algorithm is also known as *Latest Available Unscheduled Channel* (LAUC) without void filling algorithm. The intention of this approach is to minimize voids which are created when making new reservations. When there are no available wavelengths for selection by horizon scheduler, the burst are automatically dropped. Because this approach makes use of voids is easy to implement. However, this comes at a cost since it results in undesirable burst loss. To address this problem, versions of LAUC with void filling were proposed in [6] and [30] to make use of voids.

3.4.1 Latest Available Unused Channel with Void Filling (LAUC-VF)

Latest Available Unused Channel with Void Filling algorithm is similar to LAUC, the only difference between the two algorithms is that LAUC-VF keeps track of all its voids and fills them with incoming data bursts. LAUC-VF has two variations which will be discussed in the next subsections. LAUC-VF reduces voids by selecting the latest available unused data channel for each arriving data burst. According to [6], given the arrival time t of a data bursts with a duration L to the optical switching matrix, the scheduler first searches for the outgoing data channels that are available for the time period of $(t, t+L)$. The scheduler identifies the channel with the smallest gap between t and the end of the last burst packet just before t [6].

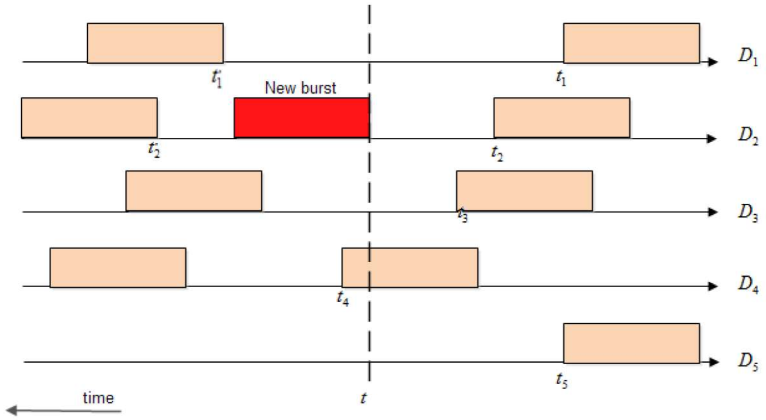

Figure 3-10: Latest Available Unused Channel with Void Filling (LAUC-VF)

An illustration of LAUC-VF algorithm is shown in Figure 3-10. From the figure, it is apparent that channel D_3 cannot be chosen because at t_t the available empty space on this channel is not enough to fit the new burst data packet. D_4 cannot not be an option because at time t , it is being used. Other channels such as D_1, D_2 and D_5 are all legible for selection. However, D_2 is selected instead of D_1 or D_5 because of the three available channels (i.e. D_1, D_2 and D_5), D_2 has the smallest gap between gap between t and the end of the last data burst just before t. That is, $t - t_2 < t - t_1 < t - t_5$. LAUC-VF algorithm is given in below.

Begin {LAUC-VF algorithm}
Step 1: i= 0; x=t ;
Step 2: j= Ch_Search(x);
if (j <> -1)
{report the selected data channel j and the selected FDL i;
stop;}
else
{
i=i+1;
if (i>B)
{report failure in finding an outgoing data channel and stop;}
else
{x=x = Q_i, goto Step 2;}
}
End {LAUC-VF algorithm}

Implementation of Min-EV has a much longer execution time than the LAUC scheduling algorithm, especially when the number of voids is significantly larger. Also it achieves a loss rate which is at least as low as LAUCVF, but can run much faster. However, it result in high bandwidth utilization and a low burst loss rate $Ch_Search(n)$is a function which searches for eligible latest available unused channel at time x, t is the data burst arrival time and j is the outgoing data channel selected to carry the data burst [6].

3.4.2 Variants of LAUC-VF

A number of the LAUC-VF algorithm variants have been proposed in the past this includes [30]; Best Fit, Minimum-Starting Void $(Min-SV)$ and Minimum-Ending Void $(Min-EV)$. Each of these algorithms has its own features that complement it and uniquely separate it from the other, for instance Min-SV is considered to be faster in implementation, and it also functions just like LAUC-VF. $(Min-EV)$ Is designed to minimize the ending void. The Best Fit ensures that the total length of the starting and the ending void generated after the reservation [5] is kept at minimum value. In Table 3-1 below a detailed comparison of the performance associated with these algorithms is given. In this case we mainly focus on the time complexity using the big-O notation, the State Information and the Bandwidth Utilization of the algorithm [5].

Table 3-1: Comparison of scheduling algorithms

Algorithms	State information	Bandwidth Utilization	Time complexity
Min-SV/EV	$S_{i,j}, E_{i,j}$	High	$O(W\, log_m)$
LAUC-VF	$S_{i,j}, E_{i,j}$	High	$O(W\, log_m)$
Best-Fit (BF)	$S_{i,j}, E_{i,j}$	High	$O(log^2 M)$
LAUC	$Horizon_i$	Low	$O(W)$

When, implementing Min-EV from the table we can observe that it consumes more time in execution as compared to the LAUC scheduling algorithm, especially when there is a significant number of voids. On the other hand Min-EV achieves a loss rate which is comparably lower than that of LAUC-VF, but LAUC-VF can execute faster. However, this results in high bandwidth utilization and low burst loss rates. BF, Min-SV and Min-EV algorithms are variants of LAUC-VF. All the void filling scheduling algorithms tend to yield better bandwidth utilization as well as burst loss rate in comparison to the LAUC algorithm. However they have corresponding longer execution times.

3.5 QoS Provisioning

The loss of data is one major concern when it comes to both circuit and packet switched networks. The main cause of data packet loss are the previously highlighted network challenges thus congestions, contention and in some cases route divergence during alternation caused by broken network links. In real-time application the opposite is true, data packet loss is not a major concern. In non-real time applications, for example, file transfer and email. This is mainly because such applications can retain or request for lost packets and these are retransmitted with little consequences given protocols that acknowledge the loss of packet retransmit are provided. The same cannot be said about real-time applications such as telemedicine, video and voice. Packet loss in such applications can have extreme consequences such as; distorted voice or unclear images. QoS is a concept of technologies tolerate applications depending on their transmission state to send and receive predictable services at different levels in terms of latency variations, delay and bandwidth utilization[33].

In this section, we briefly explore QoS provisioning in OBS networks. The main challenge in maintaining QoS provision in OBS networks is the unavailability of optical buffers/memories in the core network nodes. In OBS, the most important concern is to be able to set priorities in traffic. Thus, one should be able to uniquely identify traffics as either low or high priority, hence the need for extra QoS differentiation mechanisms can be successively implemented in OBs networks.

3.5.1 Concepts of QoS in OBS networks

There are various QoS metrics such as delay, throughput, packet delivery ratio, hop count and control overhead [34]. QoS provisioning is basically the classification of applications in terms of their priorities and functional mechanisms. Throughput is a ratio of packets sent and packets received. Delay is measured by computing the difference between the times taken to receive a packet after an initial time taken when the packet was sent. Thus QoS provisioning can either be relative or absolute. Relative QoS is the performance of a class defined with respect to other classes. Absolute QoS is not defined with any relevancy to other classes. In the next subsections we explore these categories of QoS mechanisms.

3.5.2 QoS Categories

A general overview of classification of QoS mechanisms in OBS networks for the purpose of this work only, we will focus on QoS differentiation with one-way signalling will be discussed.

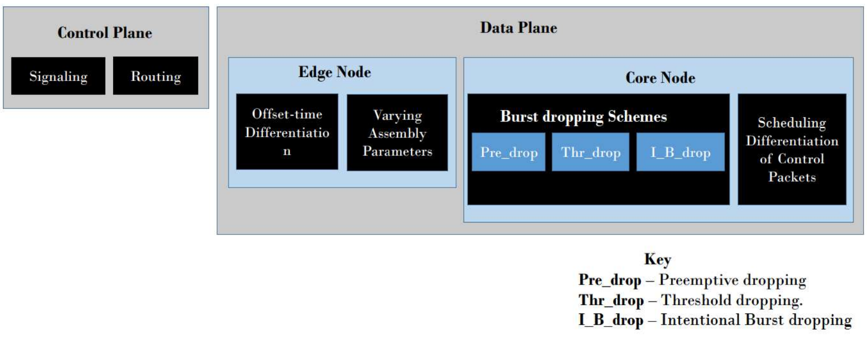

Figure 3-11: Mechanisms for QoS provisioning.

Figure 3-11 depicts the main components that can contribute to QoS provisioning in OBS networks with one-way signalling. From the illustration above these components are associated with the operation of the control plane through the signalling and routing functions and also with the data plane operation in edge nodes and core nodes. In the next subsections, we further discuss in detail the mechanisms for QoS provisioning in OBS networks. Firstly, we illustrate the categories of QoS provisioning and burst dropping as follows:

3.5.3 Edge-based QoS differentiation mechanisms

As previously discussed section 3.1.1, the purpose of the edge node is to aggregate incoming client's packets into bursts according to their destination and traffic classification. At the egress node the process of de-aggregation of bursts data payload before it is transmitted to the client network is performed. For all these a features in burst data packets a certain degree of QoS must be guaranteed. Currently there are two mechanisms that have been proposed for edge-based QoS differentiation. As illustrated by Figure 3-13, these are; varying assembly parameters $(burst-length)$ and Offset time-based differentiation.

3.5.3.1 Burst Length based Differentiation (BLD)

The idea behind BLD is to assign HP class lower timer and lower burst length thresholds than those of LP. As a result, short bursts will be able to use the voids generated by the already scheduled bursts. This will improve the performance of the HP class relative to the LP class as far as loss probabilities are concerned. According to [38], majority of traffic is assumed to be best effort, therefore to reduce the overall switching overhead in the system, low-class burst assembly threshold larger than the optical switching time must be selected. End-to-end delay is minimized by using shorter burstification timer for HP class. The disadvantage of BLD is its inherent complexity and increased signalling overheads.

3.5.3.2 Offset Time-based Differentiation

The technique involves setting extra offset-time to high priority bursts so that they can be given priority when reserving resources for them. In such a case, there is an absolute isolation between high priority $(HP)$and low priority (LP)classes. In that way, we ensure that no HP class bursts are blocked as a result of the presence of LP class bursts. This technique is renowned for its simplicity. It also reduces the loss probability of HP bursts by their postponed transmission [35]. However, its main disadvantage is that of sensitivity of high priority class to length characteristics [37], and extended pre-transmission delays which may not be tolerated by other applications.

3.5.4 Core-based QoS Differentiation Mechanisms

One of the most important functions of the core nodes is to resolve the contention problem. That is where core node based QoS provisioning takes place. This is however achieved with the help of burst dropping techniques. The three burst dropping techniques that have been proposed are; pre-emptive, threshold-based and intentional bursts dropping.

With pre-emptive dropping, when contention arises the lower priority is pre-empted by a higher priority one.

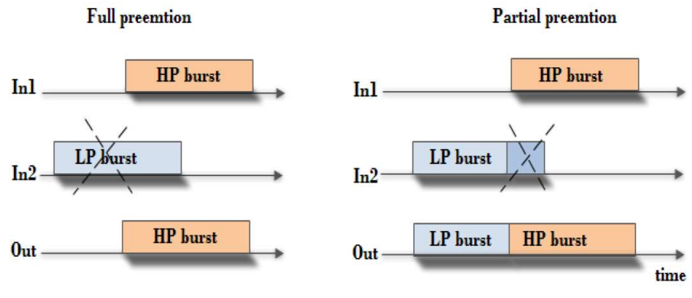

Figure 3-12: Full and partial pre-emption illustration

Thus, the LP burst is dropped / discarded. Pre-emption can either be full or partial as shown in Figure 3-12. In full pre-emption, the whole pre-empted burst is discarded while in partial pre-emption only the overlapping part of the LP reservation is affected [39]. Partial pre-emption efficiently utilizes the resources compared to full pre-emption. The disadvantage of partial pre-emption is that it imposes additional complexity on burst assembly. In threshold-based dropping mechanism more resources are given to high priority bursts than to low priority bursts. This is done in accordance to a chosen threshold parameter, for example; wavelength. Provided the required resource is below the set threshold, both HP and LP bursts will be accepted. However, as soon as the occupancy of the resource goes above the threshold, LP bursts are dropped but arriving HP bursts will be accepted provided the resources are still available. However, there is a problem associated with this mechanism. For example; when the threshold has been reached, and the resources are still available and no arrivals for HP burst, LP bursts will continue to be dropped. This leads to high drop rate for LP bursts. Intentional burst dropping is just an extension of a pre-emption based QoS differentiation scheme to provide absolute QoS. Intentional burst dropping mechanism maintains the performance objectives of HP bursts at certain levels by intentionally dropping LP bursts using an active discarding scheme [36].

3.5.5 QoS mechanisms in OBS networks with one-way signalling

Table 3-2: QoS mechanisms in OBS networks with one-way signalling

QoS mechanism	Implemented QoS model	Supported QoS parameter	Advantages	Disadvantages
Hybrid signalling	absolute	delay / burst losses	- absolute end-to-end loss and delay guarantees for HP	- lower statistical multiplexing gain, ineffient usage of bandwidth (less resources available for LP traffic)
QoS routing	absolute (delays) relative (burst losses)	delay / burst losses	- introduces QoS guarantees on network level	- controlling burst losses may be challenging (need the knowledge about network state)
Offset-time differentiation	relative	burst losses	- simple, soft operation - no need for any differentiation mechanism in core nodes	- sensitivity of HP class to burst length characteristics - extended pre-transmission delay
Varying burst assembly parameters	Absolute (delays) relative (burst losses)	delay / burst losses	- assembly parameters can be easily setup	- the resulting traffic characteristics may influence network performance
Threshold-based dropping	relative	burst losses	- Can be easily implemented	- the efficiency of bandwidth usage strongly depends on threshold adaptability to traffic changes
Intentional burst dropping	absolute	burst losses	- can provide absolute QoS	- the link utilization may suffer - complex implementation
Scheduling differentiation of control packets	relative	burst losses	- priority queuing in electrical buffers is a feasible and well-studied technique	- extended delay (need for longer queuing windows and so larger offset times to perform effectively)

3.6 Summary

In this chapter, we discussed various burst assembly algorithms. The length based algorithm performs poor in light load but better in heavier load. N_{MIN} , unlike the length based performs better in light loads and poor in heavier loads. The αL algorithm always outperforms the N_{MIN} algorithm but it is outperformed by the other algorithms. The T_{AVE} performs better than the timer and length based algorithms but it is outperformed by T_A. Of all the algorithms discussed in this chapter, the L_{MIN} , is the one which performs better [42]. However, the choice of the algorithm to use is dictated by the criterion of interest. That is, if the average burstification delay is the criterion, then L_{MIN} will be the best choice but when delay jitter is the criterion then T_{AVE} and T_A would be preferable. It is therefore desirable to have a single algorithm which will perform well in light and heavy load and also be the algorithm of choice irrespective of the criterion of interest. After the burst has been generated, resources such as output channel/wavelength need to be reserved for the outgoing burst. In the next section burst reservation protocols are briefly discussed. Furthermore, we looked into the Optical Burst Switching. Our main focus was on concepts and architectures of OBS edge node and core nodes of the OBS backbone network. This included the various functions which are performed by either the edge or the core nodes. Some of the functions are; burst assembly which deals with aggregation of packets into bursts. In burst assembly/aggregation we presented algorithms such as the Time-based, Length-threshold based and Time/Length hybrid. We also focused on burst reservation protocols, burst scheduling algorithms and contention resolution mechanisms. All these are potential areas which define the service quality in OBS network backbone channel. In this work we would like to improve these factors which affect the QoS in OBS networks. We attempt to improve on contention and congestion at all switching levels in OBS networks, by proposing an efficient buffer less technique for data transmission with minimum to zero data transmission delay. Any attempt to do so is preceded by a comprehensive study of traffic models and traffic prediction methods.

Chapter 4. Traffic Modelling

NGN will bring with it various multimedia applications with varying service requirements. Each communication service request will comprise classified burst data packet transmission handling hence the end-to-end service configuration created for each channel would be data specific for effective transmission. This makes the management of future networks to be a challenging task. Each burst data packet when being transmitted over the OBS backbone network after the configuration of the channel by the control packet tend to occupy the defined path for a certain period of time hence blocking the utilization of the channel by other communication requests. Consequently, the study of congestion and contention minimization strategy entails the identification or prediction of a best traffic model that can be implemented on any defined network to balance the channel utilization amongst communication requests and improve data transmission for all service requests. A good traffic model or analysis optimizes the way the network functions in an effort to minimize latency, maximize capacity as well as provide high reliability regardless of bandwidth available or any other network failures. On the other hand, this would absolutely give accurate estimation of the network performance based on various network variables and helps in capacity planning. Therefore, in this chapter we explore the technical operations in a bid to identify a model that we can apply in our work. Specifically, we would use it for modelling in our proposed congestion and contention minimization strategy in chapter 5.

4.1 Traffic Models

For proper resources utilization as well as dimensioning by network designers, an estimate of the traffic levels need to be made prior to any transmission. We thus overview traffic models that will assist in the design stages of such predictions. If proper models are implemented, the occurrence of both congestion and contention is likely to be minimised. Ultimately, we will select an appropriate model for applying in our work's main objective, i.e. that of minimising both congestion and contention in every backbone OBS network. The following concepts depict the available options that can be applied to this work.

4.1.1 Poisson Distribution Model

The Poisson traffic model has been utilised quite extensively over the years. The traffic model is mainly used for analysing traffic in traditional telephony networks [40] because of its buffer-less nature. The model is mainly effective if the data packet arrivals are from several independent sources, referred to as Poisson sources and will then be redistributed at random times and channel to their destination sources. The arrival packets are distributed with a parameter rate

$$\lambda : p\{A_n \leq t\} = 1 - exp(-\lambda t). \qquad (4.1)$$

Where, A_n is the traffic load, and p is probability.

In this distribution model the mean and variance are extracted from the approach used in binomial distribution, this is also widely used in most queuing models. In general a number of interesting mathematical calculations have been used in order to enhance the accuracy and feasibility of the model. For instance, the aggregation of several independent Poisson processes results in a new Poisson process. The model is also common used in traffic application scenarios where a large number of independent traffic streams is of particular importance.

Two main assumptions associated with this distribution are:-

- the number of customers/ clients is assumed to be infinitely large;
- the source data traffic arrival pattern is random;

The probability density and distribution functions for this model are:

$$f(t) = \lambda e^{-\lambda t} \qquad (4, 2)$$

$$F(t) = 1 - e^{-\lambda t} \qquad (4, 3)$$

The main conveniences associated with using Poisson distributed models is that they are manageable which makes them analytically tractable. However, Poisson models significantly underestimate the queuing performance by not taking its burstiness nature into considerations. Traditional data traffic nature in this case can be described by a random point process. This is a process whereby $\varphi = \{t_n : n \geq 1\}$ is a sequence of random points t_n at which a traffic event would have occurred, Thus;

$$0 < t_1 < t_2 < \infty \tag{4. 4}$$

With, $t_n \rightarrow \infty$ as $n \rightarrow \infty$.

The point process has two interrelated processes designed to improve its overall output generation these are: inter-arrival process and counting process. The counting process $\{N(t) : t \geq 0\}$ for φ is the number of points that fall in the interval $(0, t)$. Let $T_n = t_n - t_{n-1}$, then the process $\varphi = \{T_n : n > 0\}$ is the inter-arrival process. In this model, the inter-arrival process is exponentially distributed, thus making it memory-less, and greatly simplifying the analysis of a queuing system.

4.1.2 Pareto Distribution Process

The Pareto distribution model focuses on the time intervals between data packet transmissions over the backbone network. In its approach the idea is generate independent and identically distributed (i.i.d) data packet transmission time inter-arrival [41]. In general, if given X being random variable of any number of data packets within a Pareto distribution, the probability that X is greater than some number x :

$$P(X > x) = \left(\frac{x}{x_m}\right)^{-k}, \text{ for all } x \geq x_m \tag{4, 5}$$

The probability distribution and density functions are represented by;

$$f(t) = \beta\alpha^{\beta}t^{-\beta-1} \tag{4, 6}$$

$$F(t) = 1 - \left(\frac{\alpha}{t}\right)^{\beta},$$ (4, 7)

Where, $\alpha, \beta \geq 0 \geq \alpha$

The parameter β and α are the shape and location parameters, respectively. This model is applied to self-similar arrival in data packet traffic transmission. It is well known as the double exponential distribution. Another attribute that attached to this model is that it consists of an infinite variances when $\beta \geq 2$ and generally achieves a infinite mean, when $\beta \leq 1$

4.1.3 Autoregressive Model

Autoregressive model is a linear prediction formulae model that predicts any given output y_n of a system based on some previous record sets of outputs $\{y_k\}$ where $k > n$ and inputs x_n and $\{x_k\}$ where $k > n$. This approach is based on the number of variables of the model specification or attributes that are mathematically adjusted with minor changes for the actual prediction results. Because model is highly dependent on previously generated outputs of the system, it is therefore referred to as the Auto-Regressive Model otherwise it would be referred to as the Moving-Average-Model (MAM) if it were to be dependent on the present inputs to the system it. The are other prediction models that depend on both input and output of the system these are also known as Autoregressive-Moving Average models.

Autoregressive model of order ρ denoted as $AR(\rho)$, takes the following format:

$$X_t = R_1 X_{t-1} + R_2 X_{t-2} + \ldots + R_\rho X_{t-\rho} + W_t \quad ,$$ (4, 8)

Where, W_t, is the white noise, R_i are real numbers and X_t are the prescribed correlated random numbers. The auto-correlation function of the $AR(\rho)$ process consists of damped sine waves that are highly dependent on the state of the model solutions thus either real or imaginary. Discreet Autoregressive Model of order ρ, denoted by

$DAR(\rho)$, generates a stationery sequence of discreet random variables with a probability distribution and with an auto-correlation structure similar to that of the Autoregressive model of order ρ.

4.1.4 Weibull Distribution Process

The Weibull distribution is a versatile distribution that can take on the characteristics of the previously discussed traffic prediction distribution models, this model is based on the value of the shape parameter. It can model ON/OFF data packet sources. The distribution function in this process is given by:

$$F(t) = 1 - e^{-\left(\frac{t}{\beta}\right)^{\alpha}}, t > 0 \qquad (4, 9)$$

and the density function of this process is also given by:

$$f(t) = \alpha\beta^{-\alpha}t^{\alpha-1}e^{-\left(\frac{t}{\beta}\right)^{\alpha}}, t > 0 \qquad (4, 10)$$

where the parameters $\beta \geq 0$ and $\alpha > 0$ are the scale and location parameters.

This distribution is directly related to the normal distribution. For the density function at $\beta \leq 1$ the distribution gives an L shape in a graphical illustration and when the values of β positively increase thus $\beta > 1$ a bell shape graph is produced [42]. Weibull distribution gives a failure rate increasing with time, for $\beta > 1$. The failure rate decreases with time. At $\beta = 1$, the failure rate is constant and the lifetimes are exponentially distributed.

4.1.5 Markov – Modulated Poisson Process (MMPP)

The MMPP uses a finite number of states a traffic source on a network. It basically is a process, belonging to the class of Markov renewal processes, where arrivals occur according to a state dependent poisson process with different rates governed by a continuous-time markov chain. As the accuracy of the model increases the number of states also increases as well, this gives a linear relationship between the network states and model occurrences. Therefore, this model has correlated processes that have a tractable queuing analysis and can be used to capture the randomness at different scales. Furthermore, it is versatile to capture traffic characteristics. Assuming $M(t)$ be a Markov process with state space $\{0,1,2,\ldots,M\}$. If the probability law of a process is determined by the state of the current state of $M(t)$ completely, then the process is called a Markov modulated process. This Markov modulated process is classified by the modulated processes. A two state MMPP is shown in the diagram below together with the queuing approach.

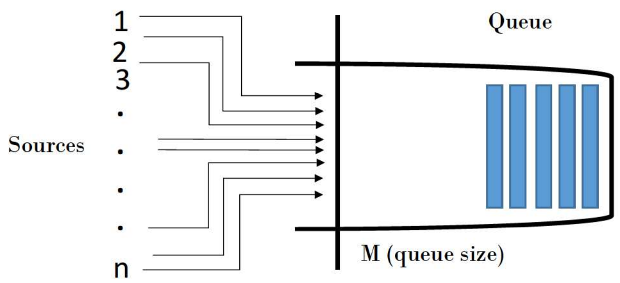

Figure 4-1: Queuing analysis for ON/OFF model

In this case the MMPP introduces some correlations between successive inter-arrival times, and therefore can capture the bursty nature of video traffic up to some extent. With different number of states of the modulating Markov chain, the MMPP can be used to model different traffic sources [43], [44]. The simplest case however, is the one that uses MMPP to model ON/OFF traffic sources.

The two state (ON and OFF) process in this model can be diagrammatized as follows:

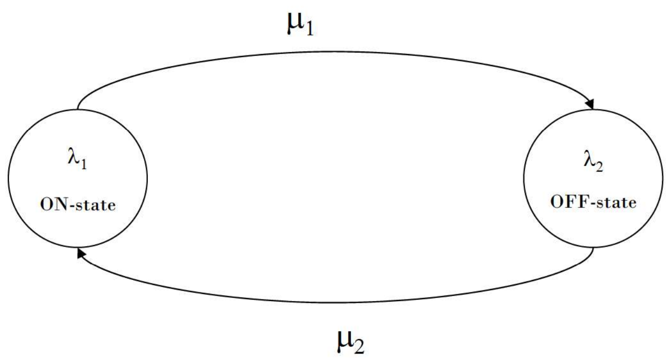

Figure 4-2: Two state (ON and OFF) MMPP

In this case the arrival during the OFF period is zero, while the arrival process during the ON period is the arrival rate of traffic. This model is sometime denoted as Interrupted Poisson Process (IPP).

4.2 Traffic Prediction

As previously indicated in the above sections, traffic modelling prediction are one of the fundamental issues in network traffic engineering and management. In designing a network for future generation applications one will be required to know the amount of incoming traffic on a particular channel/node such that the incoming burst packets would be processed according the configured QoS requirements or next hop congestion free routing channels, hence this gives us all the reasons to require a more feasible traffic predicting models in future generation networks. For the purpose of this thesis understanding traffic prediction concepts is essential because later we will propose congestion and contention minimization strategy with background alternate path searching which will be used in the eminence of congestion or contention. Hence the alternate path will use traffic prediction algorithms in order to find the best channel redirect all burst data packets. However, the main draw back in traffic predictions is that traffic patterns and behaviour change from time to time depending on the applications sending data to and from source to destinations. In section, we look at an overview of some traffic models that can be implemented in this work.

4.2.1 ARMA Predictor Models

The general model for ARMA process as illustrated in [64], the prediction is performed using previous statistical data values on each network node to predict future traffic capacity. The correlation structure in this model makes it possible to predict future traffic rate. Given that Z_t is the white noise process and B is the backward shift operator, then the general model for ARMA process is such that;

$$\phi(B)X_t = \theta(B)Z_t \tag{4.11}$$

Where B is defined as;

$BX_t = X_{t-1}$, $B^2X_t = X_{t-2}$, etc.

$\phi(B)$ and $\theta(B)$ are polynomials of B:

$$\phi(B) = 1 + \phi_1 B + \phi_2 B^2 + ... + \phi_P B^P \tag{4.12}$$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + ... + \theta_P B^P \tag{4.13}$$

The above process is $ARMA(p,q)$ process. However, when $q = 0$, $ARMA(p,q)$ becomes a $AR(p)$ process and a $MA(q)$ when $p = 0$.

4.2.2 Linear predictor- Least Mean Square Error (LMS)

LMS based traffic prediction was proposed for online real time applications such as video traffic. Two schemes in this model have been proposed time-domain Normalized Least Mean Square (NLMS) and the wavelet-domain NLMS based adaptive prediction scheme. The wavelet-domain NLMS based adaptive scheme "*exploits the redundant information in the wavelet transform coefficients for accurate prediction*" [64]. Given that $\hat{N}(n+1)$ is the estimate of the number of packet arrivals during the $(n+1)\,th$ frame, this follows that [66];

$$\hat{N}(n+1) = \sum_{i=1}^{n} w_i . N(n-j+1) \tag{4.14}$$

where w_i, $i \in 1,\ldots,N$, are the filter coefficients and h is the length of the filter. The error $e(n)$ of the ***n***-th frame is given as [61];

$$e(n) = N(n) - \hat{N}(n) \quad (4.15)$$

the coefficients of the filter are updated at each iteration as follows; [66]

$$w_i(n+1) = w_i(n) + \delta . e(n) . N(n-i+1) \quad (4.16)$$

δ is the step size.

4.2.3 FARIMA Predictor Model

FARIMA model uses time series values from current and past values. FARIMA predictor an improved version of the LRD (Long Range Dependence) process because it mainly uses the series calculation of data packets at random intervals. This illustrates the slow decrease of the autocorrelation function $\gamma(h)$ and the FARIMA process function is defined as:

$$\phi(B)(1-B)^d R_t = \theta(B) Z_t \quad (4.17)$$

where $-\frac{1}{2} < d < \frac{1}{2}$. While Z_t is white noise with zero mean and variance σ $(1-B)^d$ is the fractional difference operator.

If $-\frac{1}{2} < d < \frac{1}{2}$, R_t is a stationary process and $d = H - \frac{1}{2}$, where H is the Hurst parameter.

In [64] FARIMA (*p, d, q*) process $\{X_t : t = \ldots,-1,0,1,\ldots\}$ was defined as follows,

$$X_t = \sum_{j=0}^{\infty} \Psi_j a_{t-j} \quad (4.18)$$

and

$$a_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \quad (4.19)$$

where

$$\sum_{j=0}^{\infty} \Psi_j B^j = \Theta(B)\Phi^{-1}(B)\nabla^{-d} \tag{4.19}$$

and

$$\sum_{j=0}^{\infty} \pi_j B^j = \Phi(B)\Theta^{-1}(B)\nabla^{d} \tag{4.20}$$

letting $\hat{X}(h)$ denote the *h*-step forecast made at origin *t* of X_{t+h} at time $t+h$, then it follows that;

$$\hat{X}(h) = -\sum_{j=1}^{\infty} \pi_j^{(h)} \hat{X}_{t+h-j} \tag{4.21}$$

where

$$X_j^{(h)} = \pi_{j+h-1} - \sum_{j=1}^{h-I} \pi_i \pi_j^{(h-j)}, h > 1 \tag{4.22}$$

$\pi_j^I = \pi_j$ ***or***

$$\hat{X}_t(h) = \sum_{j=1}^{\infty} \Psi_j a_{t+h-j} \tag{4.23}$$

and the mean square error of the *h*-step forecast is given by;

$$\hat{\sigma}_t^2(h) = E(X_{t+h} - \hat{X}_t(h))^2 = \sigma^2 \sum_{j=0}^{h-1} \Psi_j^2 \ . \tag{4.24}$$

The parameters for FARIMA are estimated using historical traffic data. The factor ***d*** is calculated using the Hurst parameter *H* which is estimated using methods given in appendix A, the implementation of the wavelet method is given in appendix B. The steps to predict the future value are given by the step-by-step procedure illustrated as follows: [58]

a) Compute $y(n) = \{\theta(B)\}^{-1}\phi(B)x(n)$, $y(n)$ is FARIMA (*0, d, 0*).

b) Compute the best liner predictor of $y(n)$, by applying [64]: $\hat{y}(n) = \sum_{j=1}^{k} \beta_{kj} y(n-j)$,

(4,25)

where

$$\beta_{kj} = -\binom{k}{j} \cdot \frac{\Gamma(j-d)\Gamma(k-d-j)}{\Gamma(-d)\Gamma(k-d+1)} \cdot \tag{4.26}$$

c) Calculate the predicted value, $\hat{x}(n) = \theta\{\phi(B)\}^{-1} \hat{y}(n)$.

d) Update the previous data when actual traffic becomes available.

e) Estimate the parameters of $\phi(B)$, and $\theta(B)$.

f) Check for the convergence condition and check for all zeros of $\phi(B)$, and $\theta(B)$ to be outside the unit circle. If all conditions are met, go to step 1 and estimate the next value.

4.2.4 Linear Predictor with Dynamic Error Compensation

Linear predictor models consist of various differentiations, we have autoregressive linear predictor (ARLP), moving-average linear predictor (MALP) and autoregressive moving-average linear predictor (ARMALP) some we have briefly discussed in the previous sections. Some predictors which can also be used for traffic prediction are MSE, FARIMA and Gaussian. Most of these predictors do not take into consideration the white noise; therefore, the accuracy of such models may be questionable. It was on this drawback that motivated the evolution of linear predictor with dynamic error compensation (L-PREDEC).

Let's suppose the mean traffic data rate $(bits/ps)$ is to be calculated for interval $(t1:t2)$, $(t1:t2)$,..., where $t_{i+1} - t_i = d_t$ is the size of the observing window, then a time series of the traffic rate $R_1, R_2, R_3, \ldots, R_t$ is obtained [62]. Applying a linear predictor on the time series and based on correlations structures of the time series, a linear predictor $P_t R_{t+h}$ predicts the future value R_{t+h} of a time series $R(t)$ based on previous values of

$\boldsymbol{R_1, R_2, R_3, \ldots, R_t}$. However, in practice only limited items can be used from past history, hence;

$$P_t R_{t+h} = m + \sum_{i=1}^{n} a_i (R_{t-i+1} - m) \tag{4.27}$$

where $\boldsymbol{m}$ is the mean of the time series $R(t)$.

L-PREDEC predictor is made-up of two parts. That is; the linear prediction part;

$$LP_t R_{t+1} \tag{4.28}$$

where LP is the linear predictor and the error compensation part is given by:

$$\boldsymbol{fec(R_t - LP_{t-2} R_t)} \tag{4.29}$$

where $\boldsymbol{fec}$ is a function to calculate the error compensation from the last prediction. Therefore, combining (4.13) and (4.14) we get:

$$P_t R_{t+1} = LP_{t+1} + f_{ec}(R_t - LP_{t-1} R_t) \tag{4.30}$$

The mean of the positive bias of L-PREDEC is given by;

$$\int_{x\rangle 0} \frac{1}{\sqrt{2\pi\sigma_{err}}} \exp\left(\frac{-x^2}{2\sigma_{err}^2}\right) dx \tag{4.31}$$

where $x = f_{ec}$.

4.3 Summary

In section 4.1 we explored the various traffic models that can be implemented over the OBS network in order to manage the capacity and performance of the network. In each case of the model there are pros and cons. As we have observed in this chapter the type of network under study and the traffic properties strictly influence the selection of the traffic model to be used for analysis. Again it has also been noted that the traffic models than cannot capture or describe the statistical properties of traffic should be avoided since they consequently would give under-estimation or over-estimation of the network performance. Over and above in all kinds of networks there is no single application of a traffic model. For instance, on the models explored in this chapter, a heavy-tailed traffic would either be under-estimated or over estimated by the Poisson model and in high speed traffic networks with unexpected demand on burst packet transfers, the Pareto based traffic model would be the most ideal traffic modelling approach to be used simply because the model takes into consideration the long term correlation in packet arrival times [52]. In section 4.2 we discussed the traffic prediction models. Since the ideal target of this work is to manage the available resources, such that congestion s well as contention occurrences are minimised, we opt for Traffic predictions based models, as they would significantly improve our congestion and contention minimization strategy. As such the FARIMA traffic prediction model has proved to be a preferred candidate for data network traffic predicting since it can adequately capture self-similarity, short range and long range dependant traffic characteristics. We thus will utilise it in the evaluation of our proposed congestion and contention minimization strategy in chapter 5.

Chapter 5. Congestion and Contention Minimization

Current global data traffic is increasingly dominated by delay and loss intolerant IP traffic. This IP traffic displays a structural self-similarity. Due to the buffer less nature of optical burst switched (OBS) networks, contention and /or congestion in the core nodes can quickly lead to degradation in overall network performance at moderate to high traffic levels due to heavy burst loses. Several approaches have been explored to address this problem, notably measures that minimize burstification delays, congestion, contention and blocking at the same time enhancing end-to-end throughput as well as rational utilization of the links. In this way contention and congestion consequently lead to QoS degradation for delay as well as loss sensitive. Noting that both congestion and contention minimization are key to a consistent QoS provisioning, thus the objective in this chapter is to ultimately propose and analyse a scheme that seeks to minimise both in a bid to guarantee a consistent QoS as well as rational and fair use of available links. We address this in steps as follows: -

- First we define congestion, sources of occurrence, as well a review of possible avoidance measures and mechanisms in OBS networks.
- This is followed by contention, sources of occurrence, as well a review of possible avoidance measures and mechanisms.
- Finally, we propose and analyse a congestion and contention minimisation scheme. It is primarily a service differentiation based scheme that aims at congestion, blocking and latency minimization, by way of combining time averaged delay segmented burstification, and random shortest path selection based deflection routing and wavelength assignment.

5.1 OBS Network Congestion

Congestion as defined by the ITU-T is a state of network elements (e.g. switches, concentrators, cross-connects and transmission links) in which the network is not able to meet the negotiated network performance objectives for the already established connections and/or for any new connection requests. In general, congestion can be caused by, unpredictable statistical fluctuations of traffic flows or fault conditions within the network. In an OBS network congestion can be e.g., at link level (more traffic than what the assigned wavelength(s) can handle), at wavelength level (non- availability of a wavelength, hence continuity failure), switch level (e.g. no wavelength converters), as well as at BHCP controller level (i.e. the switch not being able to process all the incoming BHCPs in real time). Noticeable effects include burst losses loss or the blocking of new connection requests, increased queuing delays as well as end to end delays. A consequence of the burst losses or blocking of connections is the increase in the offered load ultimately leading to a drop-in network throughput.

5.1.1 Sources of Congestion

The congestion problem in OBS networks cannot be solved by merely provisioning more buffer space, higher capacity links, more wavelengths, or high speed BHCP controller processors. In essence, it is not a trivial problem because of the various requirements such as low overhead, responsiveness, fairness and so on. Location of the origins of congestion is widespread including both the edge and core. In this section, we are going to explore a few of these sources.

5.1.1.1 Edge Nodes

At the edge nodes data packets are aggregated into bursts. The burst packet generation is mainly defined by various algorithms at each node in the backbone network. From a practical point of view congestion is highly attributed by the approach in burst aggregation and burst deployment processes applied at each edge node. In the event that the configuration settings at the edge node is maintaining a small size of the burst length this implies that more data burst are going to be generated compared to a large burst length size that also implies a slight increase in delay of burst packet deployment while the edge node waits for large burst length formation. This also implies a longer offset time required for burst padding in the event that no packets have been received for data burst generation. This scenario may be graphically represented as follows:

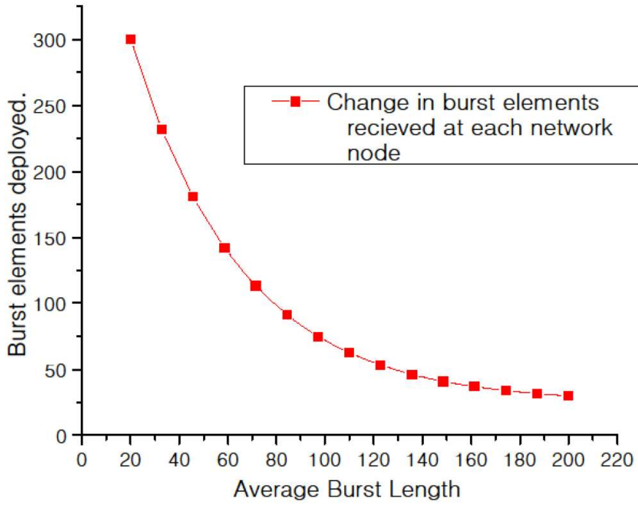

Figure 5-1: Change in Burst length maintained at the Edge node.

5.1.1.2 Burst Offset Time

In terms of offset time between the burst and its corresponding control packet, if the offset time is not adequately configured to give enough time for resources to be configured after a control packet is sent through into the network channel but before the deployment of the data burst this will generally cause congestion in the OBS backbone network. The control packet is intended to preconfigure the network resources and on arrival of the burst packet at each network node the burst packet should fly by the designated or predefined channel without any processing involved. Now if the offset time is not properly configured this will force the network switch or node to drop other connections while waiting for the current burst transmission to be completed. Only after then will the switch be available for other connections. Because of the buffer less nature of this paradigm the configurations at the switching channel are based on the information extracted from the control packet. This also follows the same criteria as the congestion at the edge node. Given enough or adequate offset time, congestion is reduced since there is no delay on the burst transmission at each network node. An increase in the offset time reduces congestion.

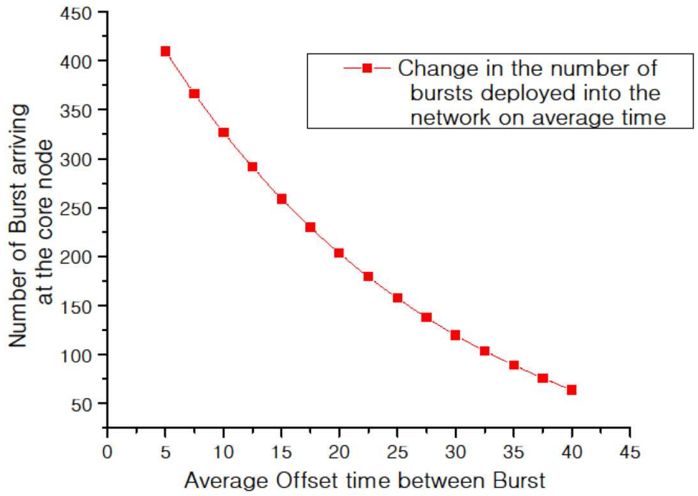

Figure 5-2: Increase or Decrease in Burst Offset time.

If the offset time is generally small, the network channels will be flooded with bursts since after a short period of every control packet deployed a burst follows. This will cause overload and consequently congestion in the network.

5.1.1.3 Burst Assembling

In burst assembling, there are various forms of algorithms that have been explored in this work. Currently the best and most effective assembling algorithm is the hybrid approach. The hybrid approach is comprised of the average time threshold algorithm and the burst length threshold. In average time the data packets are generally aggregated to form a data burst at a given average time and on the elapse of the average time the burst is compiled and deployed. In the imminence of low traffic capacity where the data is not filling the burst to the average minimum length, the data burst is filled with padding to a minimum length ($L_{\min}$) in preparation for deployment. The assembly approach can cause congestion through generating voluminous data burst that can cause congestion.

5.1.1.4 Deflection Routing

Deflection is also another source of congestion and contention. It was designed to alleviate congestion and contention conditions but as the algorithm is applied it is most likely to be a source of congestion through over utilization of the deflection routes themselves. Affected bursts designated for a particular route are redirected to an alternative route due to congestion or contention in the initially configured route. As the network traffic load increases burst are continuously deflected to a single alternative route that will in the long run be congested due to traffic increases.

5.1.2 Congestion Avoidance Schemes

Congestion control in OBS networks is quite crucial. Excessive congestion levels may lead to high burst losses. Whereas it has been studied extensively with regards to its control in ATM and IP networks, it however poses several new and unresolved scenarios for OBS networks. Since an OBS network operates largely independent of any electrical or optical buffers at its interior nodes, congestion control schemes that are applied to buffered networks like ATM differ considerably in their architecture to that of OBS networks. In an OBS network, congestion and contention are interrelated, hence their distinguishing boundaries quite ambiguous. In a way drive, each other, but congestion is relatively more transient in nature. Contention leads to burst loses due to lack of wavelengths at a given output port. But the bursts arriving immediately after the losses might just pass through fine. Congestion tends to be prolonging over a lengthier time frame and leads to increased contention problems over a longer time period

5.1.2.1 Deflection.

Since the OBS network is buffer-less in nature, deflection techniques received much attention as a key contention resolution strategy. It can work with limited optical buffering or even no buffering at all. In this case edge nodes aggregate the incoming traffic into variable length optical bursts and core nodes asynchronously switch these bursts. A key characteristic of OBS is the hybrid approach in which burst control packets are signalled out of band and processed electronically while data bursts stay in the optical domain until they reach their destination node. This avoids buffering as well as synchronization problems that are present in optical packet switching. The one-pass reservation ensures elimination of significant signalling delay. However, due to this one-pass reservation strategy and statistical multiplexing, burst loss can occur in case of contention. For this reason, deflection may be considered as an efficient resolution strategy in to achieve a low burst blocking probability. The contending bursts must be deflected to an available output port rather than delayed. This is partly illustrated in Figure 5-3. In this case this approach can work with limited optical buffering or even without buffering [69] [70]. The basic idea is to utilize other unused links in the network. Since different output links are selected for the contending bursts in a deflection routed network, contention can be avoided. However, this also has its limitations and problems. For example, if we consider a situation where a burst transmission request arrives at a given node and cannot be accommodated on any output fibre that could connect the burst to its designated next hop. The burst may be deflected from its original next hop onto a fibre connected to a different next hop node that can accommodate the burst. It then becomes the responsibility of the alternate next hop node to re-route the burst on to its destination. Sometimes it is possible that this deflected route may even pass back through the node at which the burst was originally deflected.

In general deflection proves to be a more effective strategy when implemented in a regular mash topology.

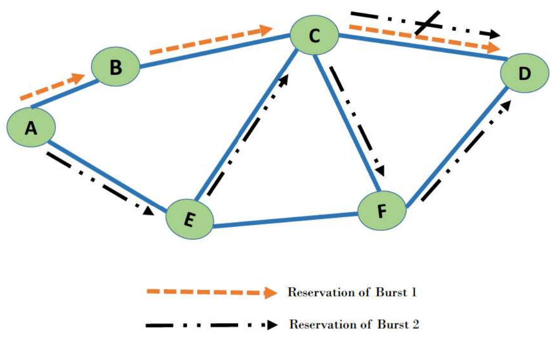

Figure 5-3: Conceptual view of deflection routing.

Figures 5-4 and 5-5 illustrate, resources reservations with and without deflection respectively.

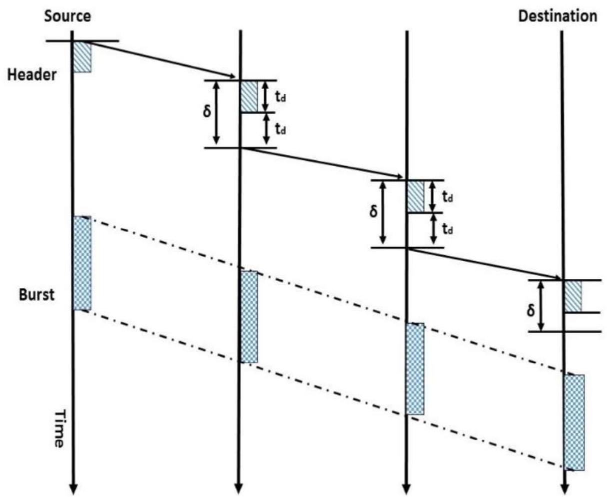

Figure 5-4: Resources Reservation Without Deflection

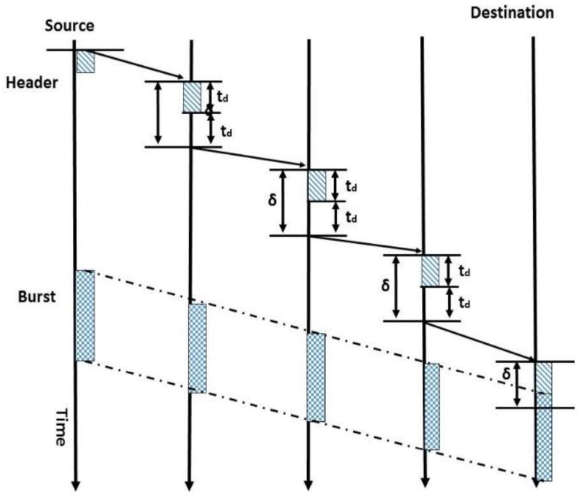

Figure 5-5: Resources Reservation with Deflection

Overall, deflection routing has proved to be a simple and yet powerful contention resolution technique. Its main drawback is that some bursts can remain in circulation within the network or rather get lost as the network traffic gradually increases.

5.1.2.2 Fibre Delay lines

In traditional packet-switched data networks, buffering is used to contain contention. Utilising the buffering approach to OBS networks is relatively complex as there is no matching RAM as yet. Instead, FDLs can be used to delay packets in the times of contention. An FDL is constructed using a piece of optic cable and thus the induced cannot be varied, as once the data packet is forwarded to the delay loop it cannot depart till it has traversed to the other end. The magnitude of delays incurable is a function of length and typically about 5μsec delay /km of fibre.

Another approach to designing a FDL buffer is by employing multiple basic delay loops of varying lengths in a parallelised structure. Depending on the type of the basic delay lines applied, FDL buffers can be classified into broad categories, namely fixed-delay and variable-delay [70]. With the fixed-delay type [70, 71], each loop provides only a prefixed delay as shown in Figure 5-6.

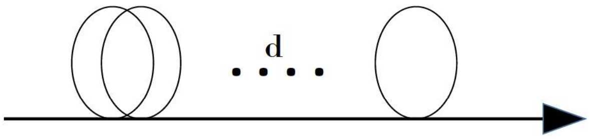

Figure 5-6: Fixed Delay-FDL basic Element.

As opposed to the Fixed Delay FDL, with variable-delay-FDL [81], each delay loop has a multistage structure, where each stage comprises 2×2 switch arrays and a fixed delay line as shown in Figure 5-7. For each stage the data burst traverses the associated delay line or the zero delay line. This makes it possible to obtain differential delays by appropriately configuring states of the switches. Such buffers can provide delay values with a much finer granularity. However they are more costly.

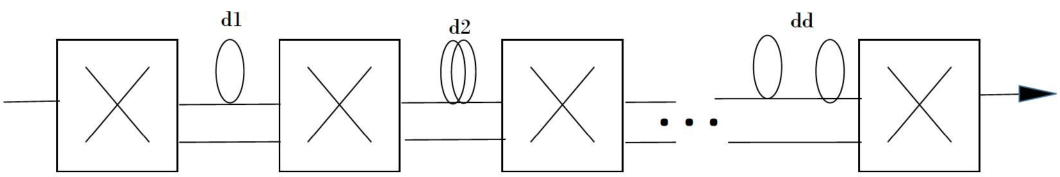

Figure 5-7: Variable Delay FDL buffer Basic Element.

It is noted that FDLs are more effective in controlling burst traffic under low to moderate traffic loads, since their increasing in number reduces the wavelength conversion requirements.

5.1.2.3 Wavelength Conversion (WC)

One of the most effective measures towards alleviating congestion is in the usage of WCs. In this case if two bursts are destined for the same output port at the same time, a WC is being used to convert one of the i/c bursts to a deferent free wavelength on the same output port in order to eliminate any possible conflicts in wavelengths.

Tuneable wavelength converters (TWC) are used to change the carrier frequency of contending bursts. If a switch can convert any input wavelength to any of the N channels at the output, it is said to have full wavelength conversion capability. Wavelength conversion can be either full or partial. In full wavelength conversion, there are as many converters as there are wavelengths, whereas in partial wavelength conversion, there are fewer converters than wavelengths. In the latter case the converters are organized in a converter pool, for which the number of converters is an essential parameter.

Wavelength conversion reduces contention thus the overall blocking probability of an optical switch is reduced. For a given normalized load, blocking probability is lower when more WCs are available. However, it should be noted that WCs are relatively expensive. In practice, there is a limit on the number of available WCs. The number of available TWCs may be a fraction of the channel count. Also, there is a limit to the tuning range of wavelength converters. A given input frequency may only be convertible to a reduced set of the output wavelengths. The effect of these constraints on the blocking probability is studied in [82] and [83].

5.1.2.4 Hybrid scheme

Dynamic resources assignment response to the resources requirements of new emerging services is a key demand as well as advantage of FG optical networks. Hybrid schemes have been devised to integrate two or more congestion and contention control measures in backbone OBS networks. In practice any of the three basic contention resolution schemes in optical networks, i.e., FDL buffering , deflection routing and wavelength conversion can bring about a reduction in contention and improving the performance, relying merely on one scheme might not be adequate to achieve an acceptable and consistent QoS and in some situations prohibitively costly.

Merely and exclusively utilising FDL buffering for contention resolution necessitates the use of a bulky fibres, which may not be cost effective. Furthermore, equipping a switch with WCs is not sufficient to alleviate the contention problem in case of an insufficiency of available channel numbers. However if the number of channels is large and the full wavelength conversion method is to be applied, it can be costly. It has been demonstrated that in certain scenarios where the network is loaded, applying deflection routing cannot improve overall performance [83].

One way of addressing the shortcomings of primary contention resolution schemes, is to deploy their hybrid derivatives in which two or more of the primary approaches are incorporated to yield a hybrid version.[82], [83], [84]. This will add degrees of flexibility in the design since the design of a hybrid contention resolution is subject to two different objective functions that require optimising:-

- The need to maximize a given performance measure such as throughput or blocking probability, with a given set of available contention resolution mechanisms.
- The overall minimization of the contention resolution resources demands subject to a projected performance measure

Some hybrid schemes may incorporate all three primary contention resolution schemes. Even though this will increase the overall complexity as well as cost of OBS network architecture. The following are perspective hybrid schemes:

- Wavelength conversion with FDL buffering (WC-FDL):
- Wavelength conversion with deflection routing (WC-DR) [119].

Wavelength conversion is common in both the above techniques for the following reasons. WCs can easily be integrated into the optical node (switch) architecture and secondly wavelength conversion schemes have relatively higher potential for improving performance. Comparing WC-FDL versus WC-DR) recent studies have shown that the earlier outperforms the latter [85]. An option of combining different types of FDL buffers and WCs has also been explored.

5.2 OBS Network Contention

The approaches discussed so are fairly classical. Burst non-discarding versions have been explored as well. There are also other techniques that have been developed specifically to address the contention problem in the OBS network. With the previous methods, if wavelength conversion, buffering or deflection routing fails, the entire burst is dropped. Another objective of the methods in this section is to reduce burst blocking probability. In contrast these schemes reduce burst blocking and promote partial burst transmission (reduce data loss). These include;

5.2.1 Sources and Causes of Contention

As discussed in the previous literature contention arises in an OBS router if two or more burst packets compete for the same output fibre on the same wavelength at the same time. A lot of techniques have been devised to address the contention problem. Some of these approaches are reactive approaches, thus they detect contention and then apply the technique and some approaches are proactive, thus they tend to avoid contention over the network nodes during normal functionality operations. Before we address the technical approaches, we will first explore the main sources of contention in OBS.

5.2.1.1 Wavelength Assignment

In optical networks wavelength assignment is designed to give a wavelength division multiplexing that definitely meets the high bandwidth requirements of next generation networks by dividing the huge transmission bandwidth channel into multiple communication channels with bandwidths of approximately 10 gigabits per second. Since the primary building block in wavelength assignment is the light path. The problem of wavelength assignment is also known as the Routing and Wavelength Assignment (RWA). This RWA addresses the importance of providing routes with non-conflicting light paths. Current approaches as discussed in the previous chapters apply static routing spectrum, incremental routing spectrum and dynamic routing spectrum. These approaches if not correctly configured they cause congestion and contention on the network. A detailed illustration of how these components cause congestion is as follows;

- **Static routing Spectrum**

In a static routing spectrum data burst are routed and allocated a light path based on the preconfigured channel for their output port in a routing table. In actual fact the entire connection set between network nodes is known. This approach causes congestion even in low traffic capacity rates, this is mainly due the pre-defined connection routes, for every control packet destined to a particular route node the network channel is never changed. It is therefore possible to have more than one data burst channel for the same output port. On arrival, all burst packets are routed to the same channel while other channels remain idle. This causes congestion in one channel. In other words, there is no flexibility of rout changing. This approach can be diagrammatized as follows;

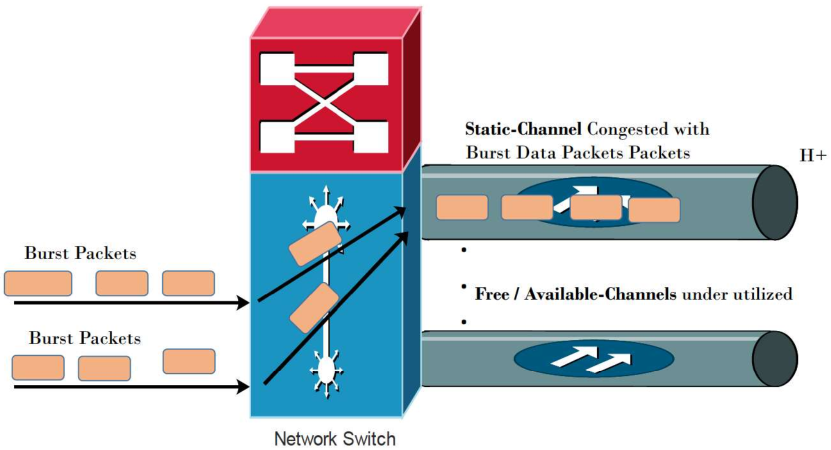

Figure 5-8: Static Routing Spectrum Incremental Routing Spectrum

In an incremental routing spectrum approach light paths are allocated to each channel sequentially. This also underutilizes resources, in most cases an initially allocated light path may be returned after a shortest transmission but that light path will not be utilized if the sequential allocation to channels has not completed the first cycle. In this approach contention is most likely to emanate in heavy traffic load since the arrival of new data burst will have to share the light paths that have not been used before because of the sequential channel allocation approach.

- **Dynamic Routing Spectrum**

This is one of the most effective wavelength assignment in this case each light path is setup for each connection request as it arrives, however the primary difference with other light paths is that, each light path is released after a finite time of utilization, this time is also known as the holding time. Hence by this approach there is a great reduction in blocking connections and maximizing the number of connections that are established in the network at any given time. Again, a poor configuration of the holding time leads to congestion and contention.

5.2.2 Contention Avoidance / Resolution schemes

Since data bursts from various source nodes may converge at a network core node randomly, contention will result if or more of them are demanding to use the same output port and wavelength at the same time. If no measures are in place to address the contention issue, all contending bursts would be discarded, this leading to a serious degradation of network performance.

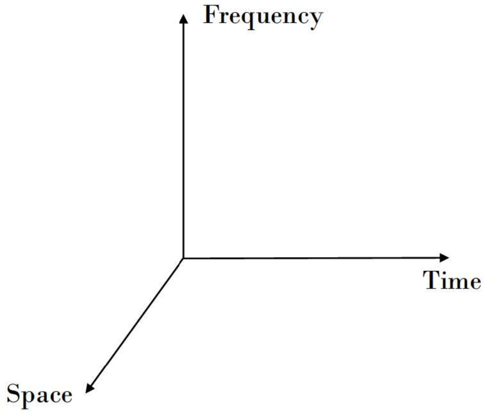

Figure 5-9. Illustration of contention control elements in the optical domain.

Because each data burst is formed as a result of aggregating several IP data packets, it is vital that contentions be managed carefully. Reactive approaches in controlling contention are predominant. Essentially, this implies the approaches attempt to resolve the contentions only after occurrence. Thus they are also referred to contention resolution methods.

As depicted in Figure. 5-9, we note three main generic directions that form the basis for resolving contention in OBS networks: time, frequency, and space, which are respectively represented by FDL buffering, wavelength conversion and de-flection routing [85]. A single contention resolution mechanism is likely to incorporate only one of these basic techniques or otherwise apply a hybrid integrated scheme that employs two or three of the techniques at the same time. Figure. 5-10 shows a taxonomy of contention resolution schemes for such networks. These are discussed further in the next section..

5.2.2.1 Contention Resolution with Wavelength Conversion

According to the commonly used Erlang's formula, a loss system with a given input traffic intensity, the probability of request blocking can be effectively lowered introducing parallel servers [86]. In fact, this approach stems from the asynchronous multiplexing gain that increases as with increase in the number of parallel servers. Usage of

WCs in WDM optical networks stems from this observation. In practice, it means a burst arriving on a given wavelength would have been blocked because the channel with the same wavelength on the appropriate output link is currently in use. However, a WC can be utilised to instead send the burst on an idle channel of the considered output link on a different wavelength that the originally designated one. Thus, the use of WCs enables inter-wavelength statistical multiplexing. Two types of WCs exists: fixed and tuneable [86], [87], [88], [89].

In fixed WCs, light paths or wavelengths at the input or output of the WC is always fixed, whereas a tuneable WC can convert any incoming wavelength to any outgoing one. There are also WCs that have a stipulated wavelength conversion range or limited wavelength conversion range, thus, for any incoming light path or wavelength it can be converted only to one of the adjacent wavelengths.

WCs can be implemented in a core switching node in various ways [88], [89]. In simplest form, they are known to perform wavelength conversion, in which one can be used per input or output wavelength channel. The figures 5-11 and 5-12 illustrate a switch architecture that incorporates full wavelength conversion at both inputs and outputs, of the switching node.

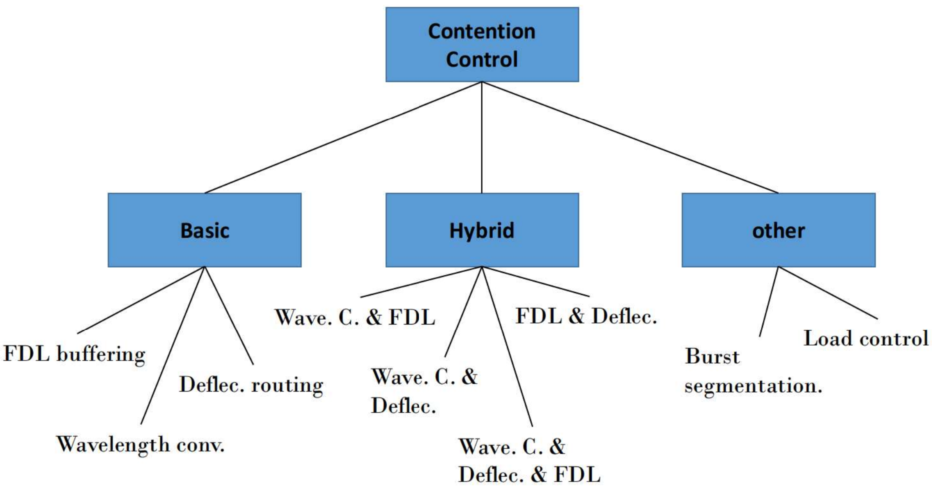

Figure 5-10: Taxonomy of classical contention control mechanisms [85].

In this case, the appropriate approach would be in the use of a fixed wavelength. For a node equipped with full wavelength conversion, a burst arrival on any wavelength can be transmitted at any wavelength at the desired output port subject to. Availing a large number of wavelength channels at any given port, full wavelength conversion can guarantee reduced blocking even though this may not be cost effective as WCs are expensive.

Alternatively, partial wavelength conversion can be a viable solution [89]. In implementing the latter, another leeway is introduced. One of the options, in this case, is to create a pool of WCs with full availability to all ports of the node, i.e., shared per node (SPN) [89] as illustrated in figure. 5-11.

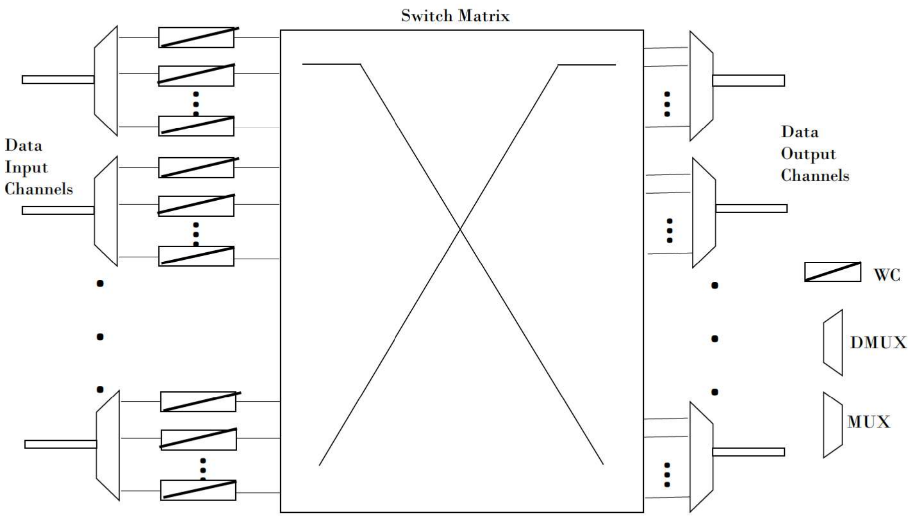

Figure 5-11 Full Wavelength Conversion at Node input.

Alternatively, another approach is to incorporate a pool of WCs at every output port [90]. Such a configuration is termed shared per link (SPL). This is illustrated in Figure. 5-14. By comparison with the full wavelength conversion, the partial wavelength conversion technique's main advantage is that of requiring a lesser number of WCs.

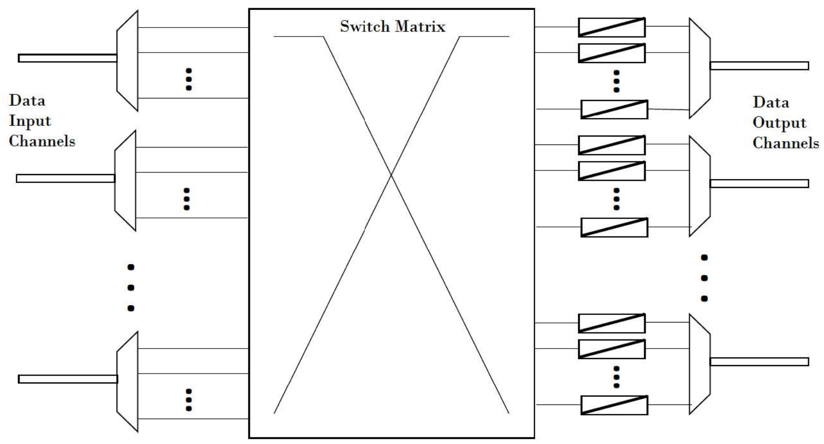

Figure 5-12 Full wavelength conversion at the Node output.

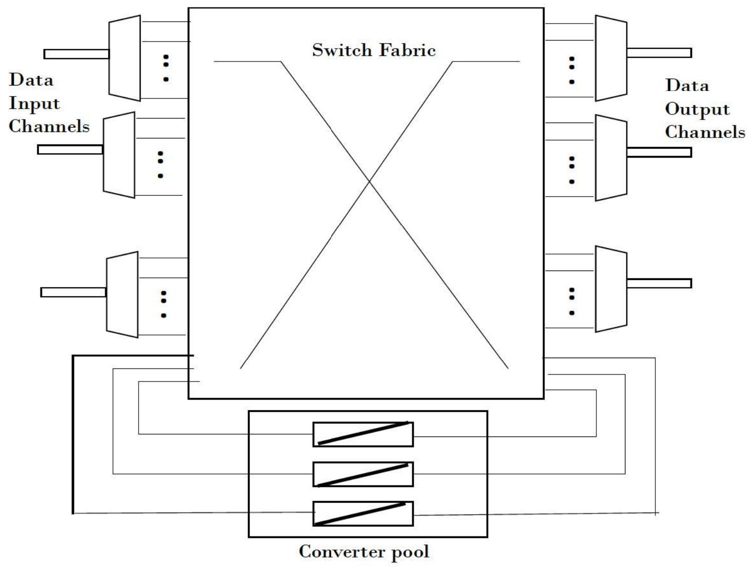

Figure 5-13. SPN Partial Wavelength Conversion.

Whereas wavelength conversion can reduce contention in OBS networks, it however will not significantly improve the overall OBS node's performance [90]. To try enhance performance further, several contention resolution schemes are concurrently incorporated. These are discussed in the next section.

5.2.2.2 Contention Resolution with FDL Buffering

In traditional packet-switched data networks, the contention problem is circumvented by buffering. The absence of optical RAM means the same approach cannot be applied directly to OBS networks. Rather, optical delay lines or sometimes referred to as loops are used to delay bursts in times of contention. A simple stretch of fibre optic cable is used to realise a delay line thus the provided delay is fixed and inflexible, since ones it is forwarded to the delay loop it cannot exit till it reaches the end of the cable stretch. The incurred lay is directly equivalent to the physical length of the loop typically 5 μsec/km of FDL.

A few graphical examples are shown next.

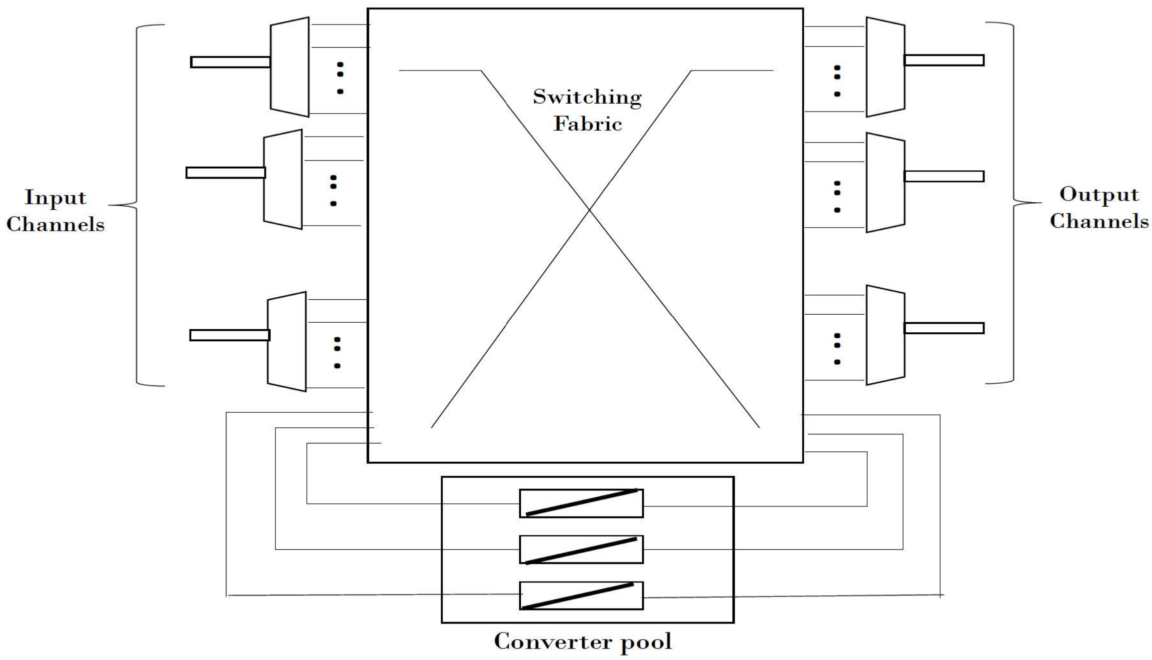

Figure 5-14 . SPL partial wavelength conversion

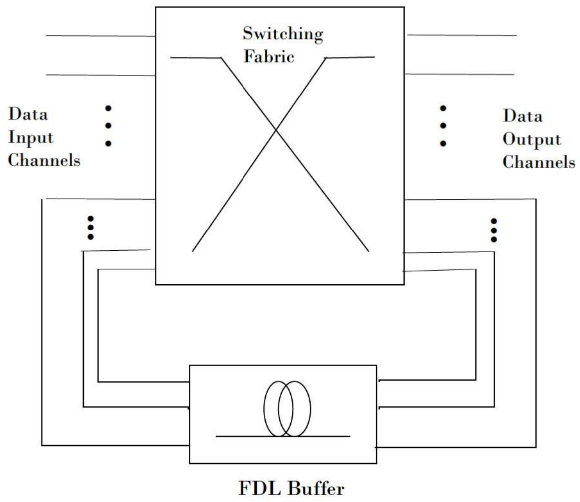

Figure 5-15 feed-back FDL buffer (single wavelength plane).

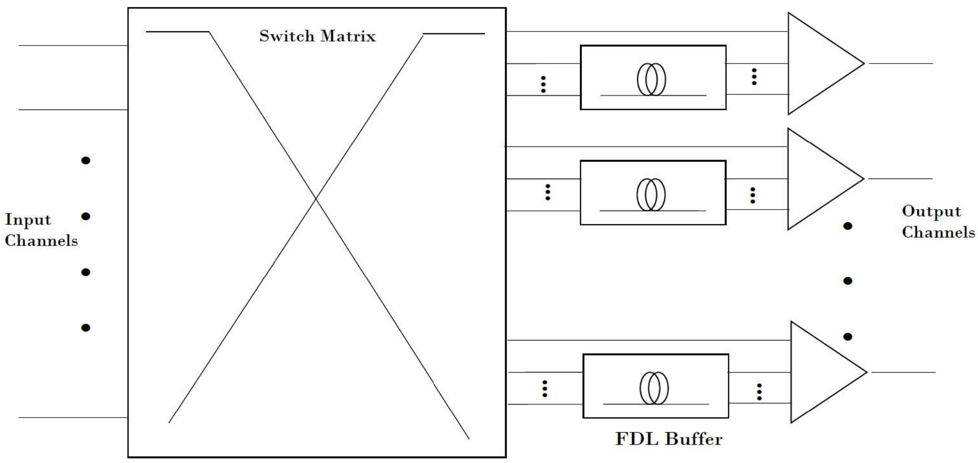

Figure 5-16   Feed-forward dedicated FDL buffer

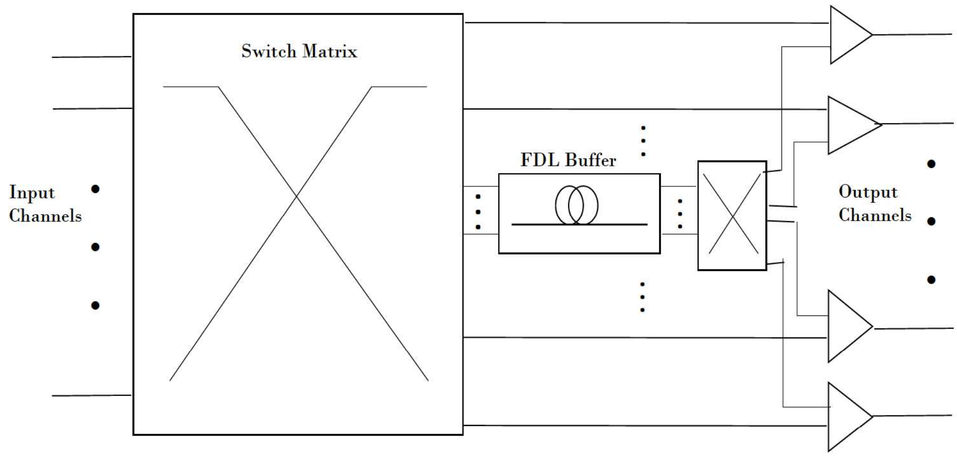

Figure 5-17 Feed-forward shared FDL buffer

5.2.2.3 Contention Resolution with Deflection Routing

Network-wide contention resolution schemes are likely to be more effective in combating contention. Such is deflection routing, which relies on the usage of all links across the network, deeming them as shared buffers. Interior switching nodes, that implement deflection routing will re-route all contending bursts on idle links, in the hope that the downstream node will redirect them to their intended destinations [89], [90].

Deflection routing is a traffic engineering technique that balances the traffic across the network. A combination of multi-path routing and load balancing techniques can be effective in equally distributing of network traffic provided the source nodes have enough information about the loading situations at various parts of the network. Accordingly information related to traffic levels across the entire network , has to be shared in real-time among all nodes of the same network, otherwise simply forwarding contending bursts to idle ports may in some situations even exasperate performance certain sections [91].

Deflection routing may also affect the performance of offset-based signaling schemes in that a burst is deflected a prior verification must be made on whether this has resulted in the increasing the overall length of the path or not. If so the offset time of the burst must be adjusted accordingly. Otherwise the burst will overtake corresponding BHP making it impossible for all downstream nodes to process the BHP beforehand. This their calls for buffering to be used long with deflection routing in case the network is using the offset-based signaling technique.

5.2.2.4 Neural Network Processing

Output contention and blocking are some of the unavoidable burst packet switching problems in OBS networks. These problems mainly emanate from the dynamic, un-scheduled burst packet arrival from the various OBS networks. In this section we will present one of the most interesting approaches in next generation networks that address the dynamic burst packet arrival [86] through simultaneous control packet processing in core network nodes. As highlighted before the contention problem is an inevitable condition. In neural network processing, the control packets contending to be routed to the same output port are distributed to other switch processors in the same switch or network node. This can be diagrammatized as follows;

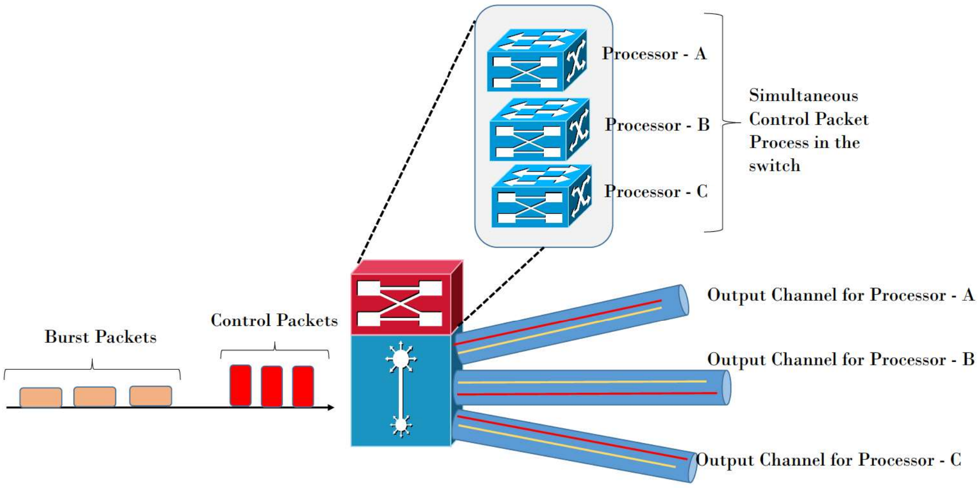

Figure 5-18: Neural Network Processor

The distribution process is also dependent on the service discipline employed by every other network node. This means the selection processes for each switching processor to be distributed the control packets will put into consideration the following criteria's; QoS demand on the control packet, FIFO selection process, longest queue selection process.

5.3 Summary of Congestion and Contention Control Mechanisms

In this work our main objective is to address the burst Congestion and contention challenges as outlined thus far as the biggest challenges of successful application of OBS network. Our work is motivated by existing literature on this topic are cantered on the reactive operations techniques after the control unit of an OBS node, detects a conflict or dysfunctional state. Therefore, in this approach usually no effort is made to avoid the occurrence of burst contentions. To understand the reason behind this trend we recall that packet-based optical switching networks, in general, and OBS, in particular, are not native approaches for optics. Actually, the basic idea of optical packet/burst switching is borrowed from the electronic packet switching, which has proven successful, and a lot of effort has been put into moulding packet switching to optics. However, some technological barriers do not allow one to one mapping of packet switching functionalities to optics, e.g. lack of optical buffers. In facing problems caused by such barriers the approaches that have usually been taken are in line with the electronic packet switching paradigm. In the specific case of packet contention, this has led to a huge amount of reactive contention resolution methods, since it is the approach that is taken in electronic networks. Nonetheless, we believe that a proactive contention control mechanism that is based on opportunities and limitations of the underlying optical network can complement the work in this field. Most of the works carried out so far in the direction of burst contention control look at the problem from a narrow perspective and consider this problem as a standalone one. For instance, in contention resolution mechanisms based on FDL buffers, the role of traffic characteristics, as shaped by the burst assembly unit, is usually overlooked and a great emphasis is put on the FDL buffer architecture.

Table 5-1: Comparison of contention and congestion resolution schemes.

Contention Resolution	**Advantages**	**Disadvantages**
Wavelength conversion	Much lower burst loss	Immature and expensive technology
FDL buffer	Conceptually simple, mature technology	Bulky FDLs; Extra delay; more voids.
Deflection	No extra hardware requirements	Out of order arrivals; possible instability
Burst segmentation	Finer contention resolution	Complicated control

The table 5-1 illustrates the advantages and disadvantages of the congestion and contention resolutions [5] as discussed in the previous sections hence giving a summary of what works best and what does not.

5.4 A Proposed Congestion and Contention Minimization Scheme

An effective approach towards alleviating congestion and contention in next generation networks should address the typical fixtures that can rectify the above-indicated sources of congestion and contention. In a brief illustration, the scheme should consist of the following desired features;

Typical features of an approach that would best address congestion and contention.

- The approach should not degrade the currently existing QoS.
- It should be applicable to all applications
- It should enhance the throughput rate as well as utilisation.
- It should ensure a rational use of resources which leads to an effective load balancing in OBS networks.

5.4.1 Model Description

In this work, we introduce and describe an approach that facilitates enhanced congestion challenges as well as congestion management thus enabling guaranteed and consistent QoS to already running applications, as well as rational and fair usage of available links. This is achieved through minimising congestion or its frequency of occurrence in the core network. In the OBS domain, congestion occurs both at edge nodes (edge congestion) as well as in the interior nodes (path congestion). Edge congestion is likely to occur when packet arrival rate is unexpectedly too high such that the edge node's limited buffering is unable to cope with the burstification process. Similarly, excessive data burst arrival rate for disassembling may aid to edge congestion. Path congestion is contributed to by contention, traffic overload as well as improper routing and wavelength assignment (RWA). The main key design and operational features of our propose scheme are summarily outlined as follows:

At the edge nodes, segmented burst assembling is assumed [92], the fundamental concept being to initially assembly data segments using traffic from multiple traffic classes with different priorities in such a way that the segments with high priority are placed towards the head of each burst and those with relatively low priority tail it. Selective deflections of sections of the composite segmented burst in case of contention (s) can be carried out, with the segments at the tail end always more likely to be dropped than those at the head end.

Figure 5-19, illustrates the burstification process, in which several delay sensitive high priority (HP) and delay insensitive low priority (LP) sources are served.

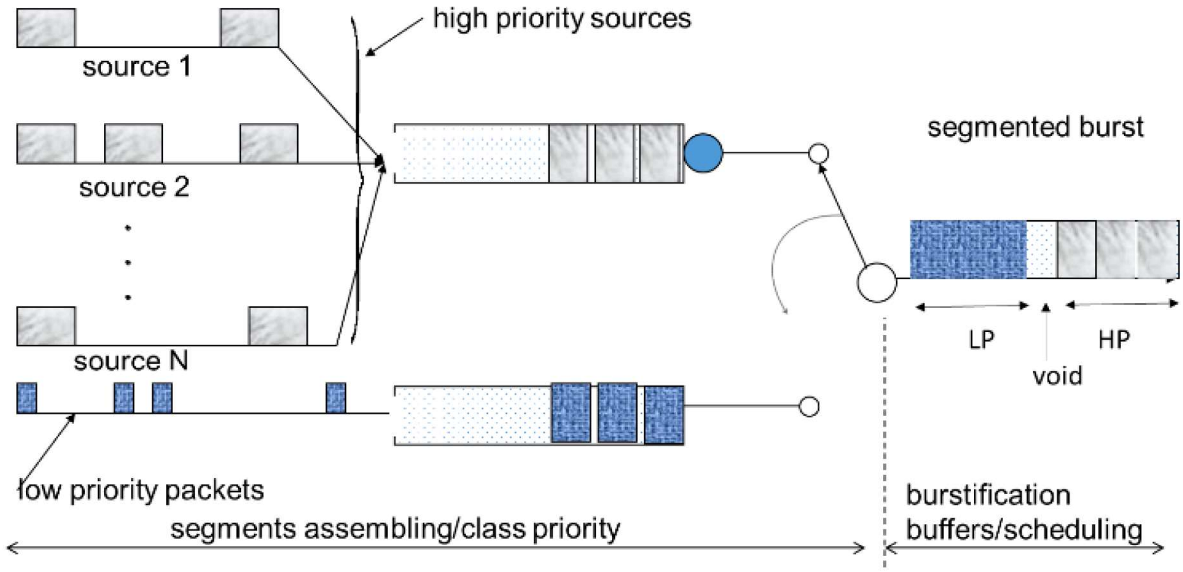

Figure 5-19: Burstification process model.

The first stage of the burstification process creates equal sized class based data segments. The assembling of all bursts is both time and length constrained. For that reason HP segments, will always be placed at the head end. This is done so as to ensure that their incurred assembling delays should not exceed a general packet average delay threshold τ , and at the same time the assembled segmented burst itself should be of minimum acceptable length $L_{min} \leq L(t) \leq L_{max}$; $L_{max} \leq 2.5L_{min}$, before dispatching. The minimum length constraint is necessary to ensure acceptable utilisation at all times.

However, in the event that T_{av} has expired, but the minimum threshold size has yet to be attained, i.e. it's still below $L_{\min}$, LP packet segments blocks, each of minimum size L_{min} are added at the tail end before the burst is dispatched. A partial void between the appended LP and HP segments allows for relative easy separation and subsequent deflection of either in the core network should contention or congestion be experienced.

Consequently, a burst is generated every τ second interval, and is common for all data flows throughout the network.

For channel reservation, we have assumed one-way reservation protocol called *Just-Enough-Time* (JET) [93], in which a data burst is transmitted after an offset time, $\Gamma_o = t_i + t_o$, seconds. This lags a control packet sent on a separate dedicated channel to reserve channels along the entire route in advance. The control packet ferries two values, offset time Γ_o, indicating expected burst arrival time, as well as burst length $L_s(t)$.

Γ_o Comprises two components, a fixed base offset time t_o which is an allowance for the control packet processing along the intended lightpath as well as a variable offset time t_i for contention/ congestion resolving if encountered along the path.

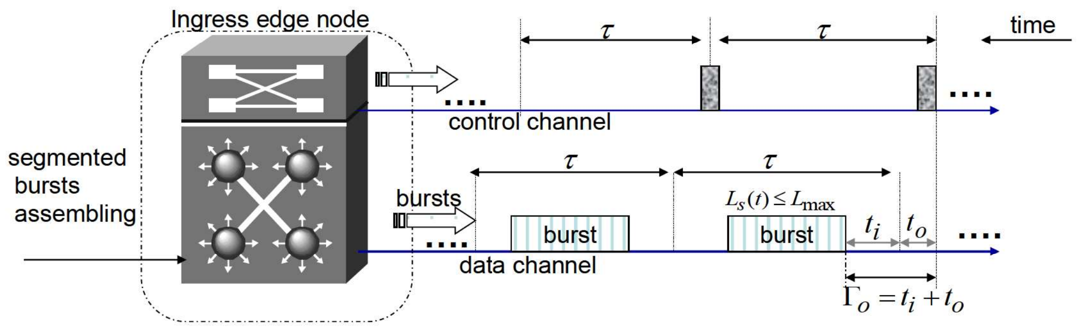

Figure 5-20: Burst assembly time.

Note that overall JET uses the offset time information transmitted in each control packet allowing to achieve higher wavelength utilization because of its delayed reservation feature. It is also fairly exigent in the information it needs to maintain.

Furthermore, JET also affords capabilities of identifying situations where no transmission conflict occurs, although the start time of a new burst may be earlier than the finishing time of an already accepted burst. The burst length information carried in the preceding control packet is used to enable close-ended rather than open-ended reservations, which would require explicit release of the configured resources. The close-ended reservation further helps OBS nodes make intelligent decisions about whether it is possible to schedule other newly arriving bursts or not.

Because, the burstification approach we have adopted inevitably generates bursts of variable sizes as well as offset times, hence may not necessarily arrive in the original send sequence as their associated control packets, and thus channel occupancy is expected to become fragmented with different voids between data bursts, hence we opt for a corresponding JET based scheduling algorithm called, *the latest available unused channel with void filling* ((LAUC-VF) scheduling algorithm since it can utilize such voids for scheduling newly arriving bursts [93]. As such it will provide a relatively better bandwidth utilization and burst loss probabilities.

With regards to the inevitable contention occurrences in the core network, which themselves are the primary cause of burst losses, we adopt a combination of deflection routing and segmentation as our contention resolution choice. Our primary goal is to segment any contending burst segment and deflect it to a non-congested route rather than discarding it. In that way, both burst loses as well as congestion are minimised.

Finally, in this approach we emphasize shortest path route selection. However, in order to reduce route computational requirements, we will use random shortest path selection. We also perform load balancing periodically so as to free up network resources at bottleneck (congested) links and in so doing support the rendering of better as well as consistent QoS end to end to users. In this regard, we take into account the “streamline effect" which is a phenomenon wherein bursts on a common link are streamlined and do not contend with each other (no intra-stream contention) until they diverge or the link confluences (merges) with other links [94]. The significance of this streamline effect is two folds: Firstly, since bursts within an input stream only contend with those from other input streams but not among themselves and their loss probabilities are lower compared with estimations using the M/M/k/k model. Secondly, the burst loss probabilities are not uniform among the input streams. The higher the burst rate of the input stream, the lower its loss probability.

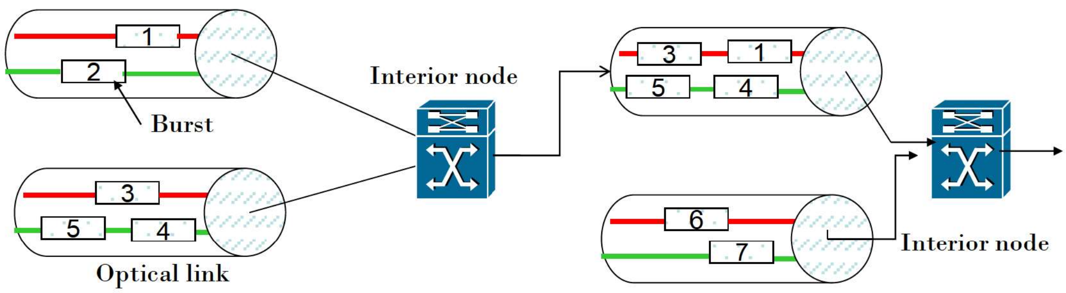

Figure 5-21: Streamline effect in OBS

This effect is illustrated in Figure 5-21. Note however that contention with other burst streams occur after merges at downstream nodes.

5.4.2 Performance Measures

In this section, we discuss some key approaches that we apply in evaluating the outlined approach. Specifically, we centre on RWA. Because edge nodes will periodically require key network status information such as, wavelength availability, link availability, traffic load and congestion in real time scales, we thus employ random search algorithms (also referred to as random search metaheuristics). It is noted that random search algorithms are considered as approximate algorithms, as they provide a near approximate solutions, but are not necessary approximation algorithms. Their key advantage is that they can provide a relatively good solution quite fast, and the same time keeping computational requirements very low quickly and easily. They also ensure convergence in probability, relatively easy to implement and comparably robust.

5.4.2.1 Random Shortest Path Searching

In this approach, all minimum hop routes between a given edge node and the desired destination node are the candidate routes. However, given the numerous links in a typical network, a key challenge is finding the distances between all pairs of communication nodes taking into metrics such as congestion, wavelength availability as well as continuity, traffic load, e.tc. In this section, we refer to these metrics collectively as weights. Thus the weight of each path can be determined by the number of nodes to be traversed, current traffic intensity on the path, as well as blocking probability etc. The entire web of nodes can be modelled as a weighted directed graph [96], [97]. Our task is to implement a search algorithm into fast compute all possible distances between all pairs of vertices(nodes) on that graph(network), with an ultimate goal of finding a set of shortest paths from any given vertex(edge node) to another (destination). Initially a matrix of all possible distances of a given graph is drawn and then after we compute its distance product [93]. If $A=(a_{ij})$ and $B=(b_{ij})$ are two $n\times n$ matrices, their distance product, $C=A*B$ is an $n\times n$ matrix such that;

$$c_{ij} = min_{k=1}^{n}\{a_{ik}+b_{kj}\} \text{ for } 1\le i, j\le n.$$

In order to construct shortest paths, the notion of *witnesses* is used. The matrix $n\times n = V$ is a witness matrix for the distance product $C=A*B$ provided for every $1\le i, j\le n$ we have $1\le v_{ij}\le m$ and $c_{ij}=a_i v_{ij}+b_{v_{ij},j}$.

We can now use a random shortest path algorithm to compute the shortest path(s) between all pairs of our network nodes, which we represent as a directed graph of n vertices in which all edge weights are taken from a set $\{-M,\ldots,0,\ldots,M\}$. The algorithm receives an $n\times n$ matrix D containing the weights (individual link states such as wavelength availability, traffic load, available QoS level, e.tc) of the edges of the directed graph (representing the OBS network) . The vertex set of the graph is $V=\{1,2,.\,.n\}$. The weight of the directed edge from i to j is d_{ij} if such an edge exists, otherwise it is $+\infty$. The algorithm commences by letting $F\leftarrow D$ before running $log_{3/2}\,n$ iterations each time setting $s \leftarrow (3/2)^{\ell}$. It then utilises a function called **rand** to generates a random subset B of $V = \{1,2,...n\}$ which is a matrix whose individual elements are selected independently with a probability $p = (9 \ln n)/s$ .If $p \geq 1$, **rand** returns the set V . Matrices $F[*.B]$ and $F[B,*]$ where $F[\cdot.B]$ represents the matrix whose columns are the columns of F that correspond to the vertices of B, whilst $F[B,*]$ is a matrix whose rows are the rows of F that correspond to the vertices of B. The distance product of matrices $F[*.B]$and $F[B,*]$ is finally computed at the same time putting a limit sM on the absolute values of all entries. Finally a comparison of F' versus F is carried out. When an entry in F' is smaller, it is copied to F as well as its corresponding witness copied to V . The search algorithm is summarized as follows [13]:

Algorithm I

$F \leftarrow D \; : W \leftarrow 0$

$M \leftarrow \max\{|d_{ij}| : d_{ij} \neq +\infty\}$

for $\ell \leftarrow 1$ to $\lfloor log_{3/2}\, n \rfloor$ do

begin

$s \leftarrow (3/2)^{\ell}$

$B \leftarrow$ **rand**$(\{1,2,\ldots,n\}, (9 \ln n)/s)$

$(F', W_{\cdot}) \leftarrow dist - prod(F[*,B], F[B,*], sM)$

for every $1 \leq i, j \leq n$ do

if $f'_{ij} < f_{ij}$ then $f_{ij} \leftarrow f'_{ij}; v_{ij} \leftarrow b_{\omega'_{ij}}$ end if

end

return (F, W)

Our next task is to implement the random path algorithm, in conjunction with wavelength assignment, with the overall objective being to maximise the number of lightpaths (N_{sd}) demands in the network, from a limited number, N_{λ} available. We define a connected undirected graph $G = (V, E)$, of a network consisting of V nodes and E edge nodes [14]. Each edge represents a pair of unidirectional fiber links in opposite directions. Thus a request for a lightpath from node i to j does not imply a lightpath in the reverse direction automatically. We let $g_{ij} = N \times N$ be the matrix such that $g_{ij} = 1$, if there exists an edge between the nodes i and j or else $g_{ij} = 0$, and also let $V = (v_i) = N \times N = 1$, if the node has been visited already, or 0 otherwise.

Our objective is to maximise the number of lightpaths than can be established for a given number of wavelength N_{λ} . i.e., maximise: $F = \sum_{i=1}^{N_{\lambda}} \varphi_i$ where, the function φ_i is defined as follows:

Let N_{lp}^{G} be the total number of light paths in the network G and letting d_{ij} be the $N_{lp}^{G} \times N_{lp}^{G}$ such that:

$d_{ij} = 1$, Only when lightpaths i and j share an optical link, and 0 otherwise.

We further define the following:

$N_{s,d}^{\lambda}$, is the total number of source-/destination lightpath demands.

$\lfloor N_{lp}(i,j) \rfloor$, to be the set of lightpaths from i and j.

$\lfloor N_{lp}^{T}(i,j) \rfloor$, be the total set of lightpaths from i to j.

$\lfloor PL_i \rfloor = 1 \times \lfloor N_{lp}^{T}(l,j) \rfloor = lightpaths(s,d),\ \forall \{s,d\} \in \{1,2,..N\}$. $\lfloor PL^{G} \rfloor = \lfloor PL^{G_i} \rfloor = 1 \times \lfloor N_{lp}^{T}(i,j) \rfloor$ vector such that

$\lfloor PL^{G_i} \rfloor = lightpaths(G) = PL_1 \cup PL_{2,\ldots PL_{N_{s,d}^{\lambda}}}$

$[P]=[P_i]=1\times N_{lp}^G$ **vector such that:**

$$[P_i]=\begin{cases}\lambda, \forall\lambda\in\{1,2,\ldots,N_{lp}^G\} \text{if the the } \lambda \text{ is assigned to lightpath } i\\ 0, \text{ otherwise}\end{cases} \qquad \varphi_i=\begin{cases}1, \text{ if } P_i\neq 0 \forall i\in\{1,2,\ldots,N_{lp}^G\}\\ 0, \text{ otherwise}\end{cases}$$

If

$$d_{ij}=1, \; \boldsymbol{i\neq j} \Rightarrow P_i\neq P_j \;\; \forall i,j\in\{1,2,3,..N_{lp}^G\}$$

$\sum_{j=1}^{N_{lp}^G}\varphi_i\leq\xi, \forall \boldsymbol{i}\in\{1,2,3,\ldots N_{lp}^G\}$, **where** ξ **is the aggregate traffic representing all demands.**

Finally, we summarise the random shortest paths election as follows:

Algorithm II

%. **Initialisation,** $k=0;\ \boldsymbol{P}=[\ ]\ \boldsymbol{P}_{max}=[\ ];$

set cycle : **for** $i=1,$ to N_{lp}^G **do**

$\boldsymbol{P}[i]=0, \&\& \boldsymbol{P}_{max}[i]=0;$

end for

$\boldsymbol{F}_{max} == \boldsymbol{base\_value};$

while $k<max\_interations$ **do**

% λ **assigning subject to constraints**

$\boldsymbol{P} := \boldsymbol{random\_solution}$

$\boldsymbol{Evaluate\_F\_of\_P}$

% **Updating** $\boldsymbol{P}_{max}$

if $d_{ij}=1, \; \boldsymbol{i\neq j} \Rightarrow P_i\neq P_j \;\; \forall i,j\in\{1,2,3,..N_{lp}^G\}$ **&&** $\boldsymbol{F}>\boldsymbol{F}_{max};$

$F\rightarrow F_{\max};$

for $\boldsymbol{i}=1,$ to $N_{lp}^G;$

$\boldsymbol{P}_{max}[\ ]\rightarrow\boldsymbol{P}[i];$

end_ for

end _if

$\boldsymbol{k}=\boldsymbol{k}+1;$

end while

% copying back

for $i = 1,$ to N_{lp}^{G}; **do**

$P[i] \rightarrow P_{max}[\ ]$

end_for

5.4.2.2 Contention loss minimization

Techniques such as FDL buffering, deflection routing, wavelength conversion, optical re-transmission of busts, segmentation and selective discarding are primarily contention resolution reactive measures. Proactive schemes would generally lead to contention avoidance and consequently congestion minimisation. Examples of such schemes include, coordinated packet transmission, in which each edge switch outputs bursts on a given set of wavelengths such that aggregated bursts arrival rate to the next core switch is kept within a prescribed threshold. Such a scheme is only effective at relatively low traffic loads.

Symmetric traffic transmission is another technique in which the edge switch sends its traffic in a symmetric manner to different output fibre links of the interior switch to which it is connected, and in that way the burst loss ratio is minimized. Load-balanced traffic in the interior network is also geared towards balancing the loads on the individual wavelengths and is achieved partly by monitoring both congestion and traffic loads, as well as their forecasting and availing this key information to all edge nodes to aid them in their routing decisions, so that contention is not exasperated. In our proposed scheme, we resolve contention by way of deflecting some portions of the segmented bursts to alternate routes. This is illustrated in Figure 5-22. As an example in Figure 5-22(i), a 50% overlapping is considered, i.e. the LP segments of bursts a' overlaps with the HP segments of burst b thus the earlier is truncated and deflected to an alternate route.

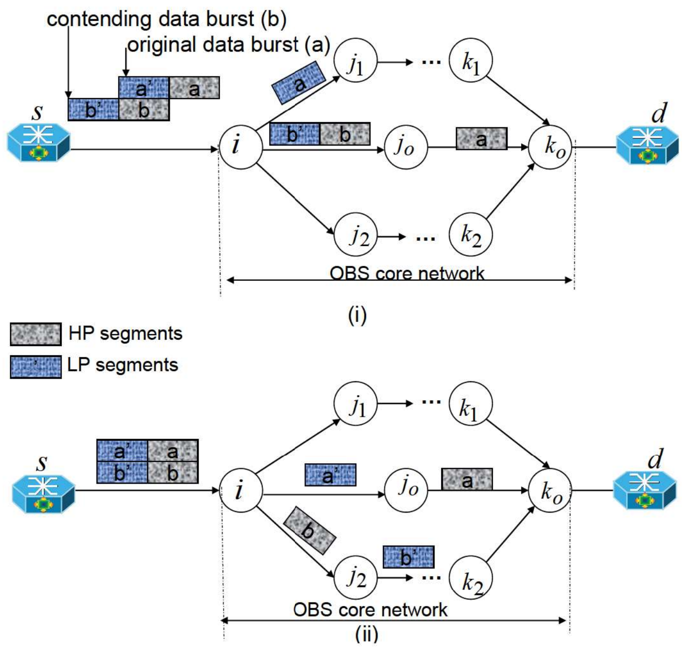

Figure 5-22: Contention resolution

In Figure 5-22(ii), a near to 100% overlapping will lead to one of the HP sections being deflected. Any further deadlock is resolved using the least remaining hops criteria [87].

5.4.2.3 Resources Allocation

Because the our scheme's objective is to minimize as well as avoid both congestion and contention, it is therefore an important design consideration that the alternate route choices be carried out taking into consideration current network conditions (link weights) e.g. traffic load, congestion status as well as utilisation link/ wavelength utilization.

Routing and Wavelength Assignment: We start by considering an OBS network comprising V nodes and L links, such that it is modelled as a unidirectional graph $G(V,L)$. Each link $l_{(i,j)}$ has a maximum capacity $u_{i,j}$, and if the total flow from a source node s to a destination node d is Φ, then our goal is to maximise:

$$\sum_{i,j\in V}(\Phi) \tag{5.1}$$

For each node $k, k \in V$ in the network, the aggregate incoming flows is;

$$\sum_{i:(i,k)\in L} \Phi'_{i,k} \tag{5.2}$$

whereas aggregate outgoing flows is;

$$\sum_{j:(j,k)\in L} \Phi'_{k,j} \tag{5.3}$$

By letting b_k be the total demand at each node k, we have:

$$\sum_{i:(i,k)\in L} \Phi'_{i,k} - \sum_{j:(k,j)\in L} \Phi'_{j,k} = -b_k, \ \ k \in V \tag{5.4}$$

and for $\Phi'_{i,j} \geq 0$ and $(i,j) \in L$.

$$\sum_{i:(i,k)\in L} \Phi'_{i,k} - \sum_{j:(k,j)\in L} \Phi'_{j,k} = \begin{cases} \Phi, & \text{if } k = s \\ -\Phi & \text{if } k = d, k \in V \\ 0, & \text{otherwise} \end{cases} \tag{5.5}$$

Similarly, the maximum throughput will be achieved by minimising;

$$\sum_{s,d\in V} \Psi_{s,d}, \tag{5.6}$$

Where Ψ_{sd} is the amount of unsupported traffic. Note that in OBS networks, the wavelength constraint has to be taken into account as well. In this regard, we first re-define a traffic matrix D representing the required bandwidth between $s, d \in V$ .the number of wavelengths is also fixed, $N_\lambda, \lambda \in N_\lambda$. We thus have an objective function to minimise;

$$min \sum_{(s,d)\in V^2 : s\neq d} \sum_{l=L} \Phi^{s,d}_{l,\lambda} + \theta \sum_{(s,d)\in V^2 : s\neq d} \Phi^{e}_{l,\lambda} \tag{5.7}$$

where $\Phi^{s,d}_{l,\lambda}$, denotes non-availability of required resources (i.e. it is equal to 1 if the required bandwidth and wavelength are available, or 0 otherwise).

$\Phi^{e}_{l,\lambda}$ reflects the quantity of traffic that could not be served because of unsatisfied constraints. θ is an objective parameter such that for all $v, s, d \in V$

$$\sum_{l\in\xi^{++}(v)} \phi_l^{s,d} - \sum_{l\in\xi^{-}(v)} \phi_l^{s,d} = \begin{cases} +D_{s,d} - \Phi^{e}_{l,\lambda} & \text{if } v = s \\ -D_{s,d} + \Phi^{e}_{l,\lambda} & \text{if } v = d \\ 0, & \text{otherwise} \end{cases} \tag{5.8}$$

In the above $\xi^{+}(v)$ and $\xi^{-}(v)$ are respectively the set of outgoing and incoming links at node $v \in V$

$$\sum_{s,d\in V} \phi_l^{s,d} \le \sum_{l,\lambda} \Phi^{s,d}_{l,\lambda}, \mathrm{c} \quad \lambda \in N_{\lambda}, l \in L \tag{5.9}$$

$$\sum_{s,d\in V} \phi_l^{s,d} \le C, \ l \in L \tag{5.10}$$

$$\sum_{l,\lambda} \Phi^{s,d}_{l,\lambda} < 1, \quad \lambda \ \in N_{\lambda}, l \in L \tag{5.11}$$

$$\phi_l^{s,d} \ge 0, \ \Phi^{s,d}_{l,\lambda} \ge 0 \ \ \{s,d\} \in V, \lambda \ \in N_{\lambda}, l \in L \tag{5.12}$$

$$\Phi^{s,d}_{l,\lambda} \in \{0,1\}, \quad \{s,d\} \in V, \lambda \ \in N_{\lambda}, l \in L \tag{5.13}$$

where the constant C represents, the capacity of each wavelength. In the formulation above we have so far assumed the Erlang B formula (applied to a single traffic tributary) as well as the sub-wavelength granularity in OBS. A good design strategy therefore would be to ensure maximisation of the occupancy of each already merged flow, by minimizing merging of traffic flows.

By taking into account the SLE [94], the traffic intensity, $\rho = \Phi_k / \mu$ is re-defined as the ratio of , Φ_k , the magnitude of serviceable aggregated traffic from various traffic tributaries at node k and μ , a measure of the quantity of traffic the node can transmit.

$$\sum_{s,d\in V}\phi_l^{s,d} \le \sum_{l,\lambda}\Phi_{l,\lambda}^{s,d}, C.K \quad \lambda \in N_\lambda, l \in L_{utilised} \tag{5.14}$$

where, $L_{utilised}$ is the number of outgoing links of the merging node v, and K is a threshold constant that ensures that the egress link is used efficiently. Optimal performance is achieved when $C.K \to \rho_{max}$ in which case contention/congestion are minimal and hence no burst losses are experienced.

Utilisation: To further dimension the available network resources optimally, and effectively it is necessary to take into account both the prevailing wavelength as well as link utilisation of each chosen link. Wavelength utilisation is the fractional time that a particular link is utilised, whereas link utilisation would be the fractional proportion of all available wavelengths in a given link path in use. Overall the utilisation $U_{s,d}$, (along a selected path from source s to destination d consists of both effective U_{eff} and ineffective U_{ineff} utilization. The former accounts for wavelengths/links that successfully reach the destination, whereas the latter accounts for bursts that are discarded before reaching their destination. U_{ineff} Tends to dominate in periods of heavy traffic. Generally, network designers ought to dimension the network resources such that both link/path and wavelength effective utilisations are rationally and fairly maximised, thus U_{ineff} must always be minimised.

In order to accomplish that, at each core node the following related information is availed.

$l_{(i,j)}$ - Possible links between two nodes along the route S to d .

τ: - fixed interval over which the load of the links are calculated.

succ(l,τ): - set of bursts that successfully traversed link l in a given set interval τ .

T_i :- length of burst i .

From the acquired above, the effective utilization of that particular link $U_{eff}(l,\tau)$ would be the fractional total occupational time:

$$U_{eff}(l,\tau) = \frac{\sum_{i\in} succ(l,t)T_i}{N_{l,\lambda}\tau} \tag{5.15}$$

where $N_{l,\lambda}$ is the total number of wavelengths comprising that link. We assume that $U_{eff}(l,0)=0$.

By letting $\{l_k, k = 1,\dots,m|\pi_z|\}$ be the set of possible paths π_z, $\pi_z, z = 1,\dots,m$, from s to d. In this case, at time t the chosen alternate path is $\pi_{z*(t)}$ whose index is obtained from the following formula [90].

$$z*(\tau) = arg \max_{1\le z\le mx} \frac{1 - max_{1\le k\le k\le|\pi_z} U_{eff}(l_k,\tau)}{|\pi_z|} \tag{5.16}$$

Weighted Least Load Criteria: If we let ρ_l^{MAX} be the maximum load threshold, which when exceeded will cause either congestion or contention on the link; then for $\rho(i,j) \ge \rho^{MAX}$ the link is overloaded hence will not be chosen. So only the alternate path with $min\,\rho(i,j)_{\in(s,d)} < \rho^{MAX}$ will be chosen.

Weighted least congestion & distance Criteria: The primary objective of this criteria is to deflect and re-route the truncated bursts on the least congested route as well as least number of hops. If we define the load on a single selected link as:

$$p_{i,j}(\tau) = \frac{N_{sdrop}(l_{i,j},\tau)}{N_{drop}(l_{i,j},\tau) + N_{succ}(l_{i,j},\tau)} \tag{5.17}$$

where,

$N_{drop}(l_{i,j},\tau)$: - is the number of dropped bursts/segments on link $l_{i,j}$ over the observation interval τ .

$N_{succ}(l_{i,j},\tau)$: - is the number of successfully transmitted bursts on link $l_{i,j}$ over the same observation interval τ .

The route choice selection can now be based on a weighted function based on congestion as well as extra distance.

$$W_{(i,j)} = \rho_{i,j}(\tau) + \frac{d_{i,j}}{d^{MAX}} \quad (5.18)$$

5.4.3 Performance Evaluation

In this section the effectiveness of the proposed scheme is evaluated through simulating on a 10 node network using OMNET++ with the necessary OBS modules and algorithms incorporated into it. Only 3 of the edge nodes are designated ingress, whereas the rest are core.

We make several assumptions, pertaining to this simulation evaluation as follows: -.

- At the ingress nodes, all traffic sources are categorized into either HP (delay sensitive) or LP (delay insensitive), and that segmented burstification is assumed.
- That the links throughout the network are of varying lengths as indicated against each in Figure 5-23.
- Each link, $l(i,j)$ is bidirectional i.e., comprises two fibers in either direction. In turn, each fiber has 32 wavelengths, 4 of which are reserved for control purposes.

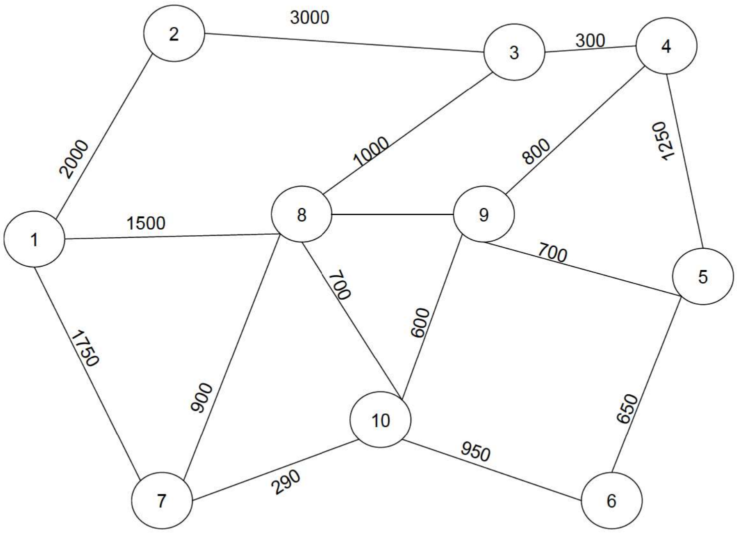

Figure 5-23: Ten (10) Node model

The maximum data rate on each wavelength is limited to $10Gbps$.

- Control Packets processing and switching fabric configuring time (switching latency) in the core switches is $1.5\mu s$.
- JET, one way signalling and LAUC-VF scheduling are assumed.
- Random shortest path route selection using a modified version of the algorithm in [98] is used to find the path between all node pairs, subject to all other RWA constraints being met.
- Physical impairments and their effects are not taken into account.
- Each node handles both transit and locally generated traffic flows, in which case the SLE is taken into account.
- Network wise, all bursts are assumed to be uniformly distributed over all sender-receiver pairs.
- There is no buffering in the core nodes and as such contention resolution is by way of segmenting the contending (overlapping) sections and re-routing on alternate paths.

- The least remaining hops policy is implemented during contention resolution.

We have chosen segmented burstification, as outlined earlier, so as to facilitate easier separation of tail end LP sections when contention occurs in the interior nodes.

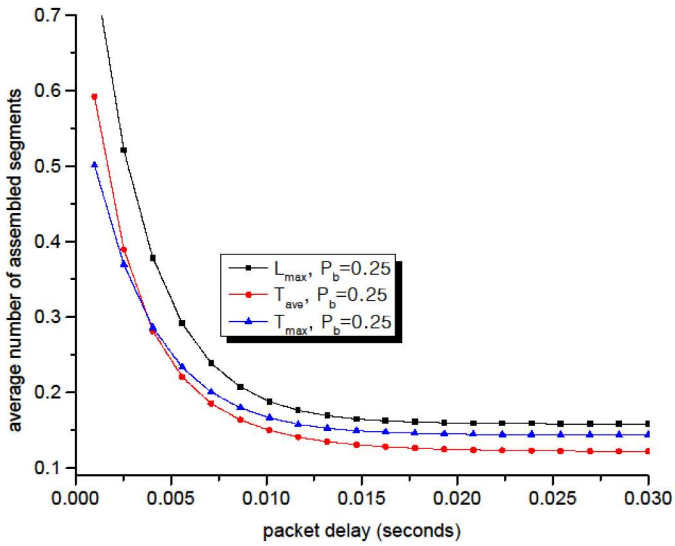

Figure 5-24: Number of segments versus packet delay.

Segment size is $30kB$, $\rho = 0.25$

In Figure 5-24, the average number of data segments generated is plotted against packet delay for the, L_{max}, T_{max} and T_{ave} burstification algorithms. The primary goal here is minimise burst assembly delays at the same time, ensuring that the burst segments are of acceptable minimum sizes. It is observed that at approximately $0.003secs$, the T_{ave} scheme an acceptable average number of burst segments equalling 0.3. This is the same with that of the T_{max} scheme. In that the QoS of applications is not compromised due to excessive delays, neither is the segment sizes too small.

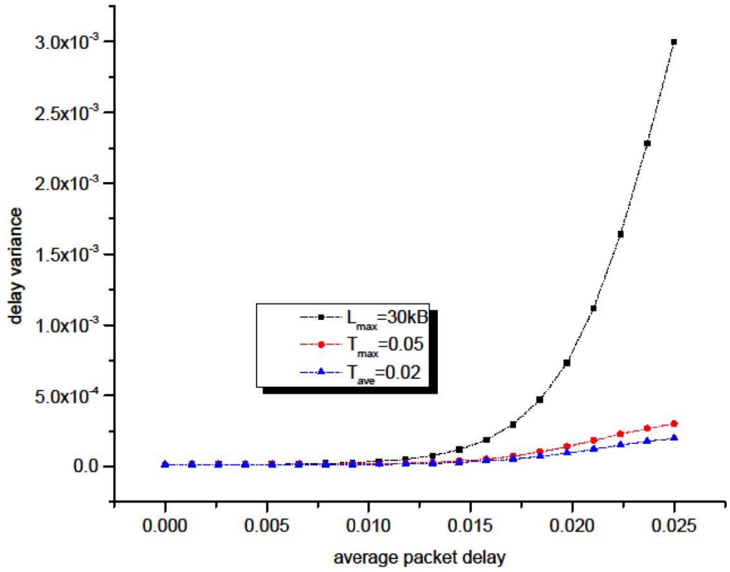

Figure 5-25: packed delay variance versus average

Packet delays for $\rho = 0.35$. The segment assembling queues (Figure 5-25) are fed by a Poisson-Paretto packet generating modelled source. It can be observed that for a given average packet delay, L_{max} generates more segments, than the other two algorithms (T_{max} and T_{ave}). Consequently it implies that the smaller the segment size, the better performing is the assembly algorithm, as the individual packets will incur lesser delays. This enhances QoS for delay sensitive (HP) traffic.

Figure 5-25 shows the variance of the packet delay comparisons for the three segment assembly algorithms. $L_{\max}$ comparatively performs much worse as the variance increases exponentially with increases in average packet delays Note that for fairly moderate to heavy traffic loads, the delay variances for both the, T_{ave} and T_{max} have a tendency to converge, since heavy packet generation rates lead to uniform packet delay.

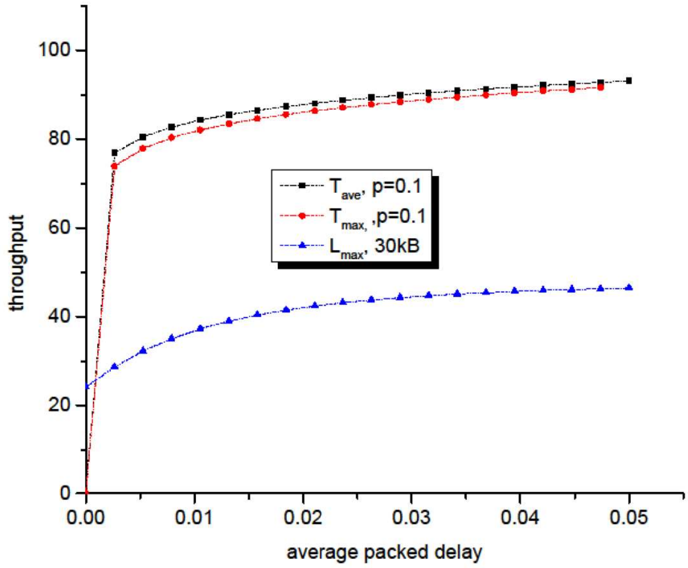

Figure 5-26: Throughput versus average packet delay.

A plot of the throughput γ versus the average packed delay is shown in Figure 5-26, in which we observe that the T_{ave} scheme outperforms the T_{max} scheme, when large average packet delays can be tolerated. Figure 5-27 shows the average number of packets per segment versus the average delay allowed for the assembly process.

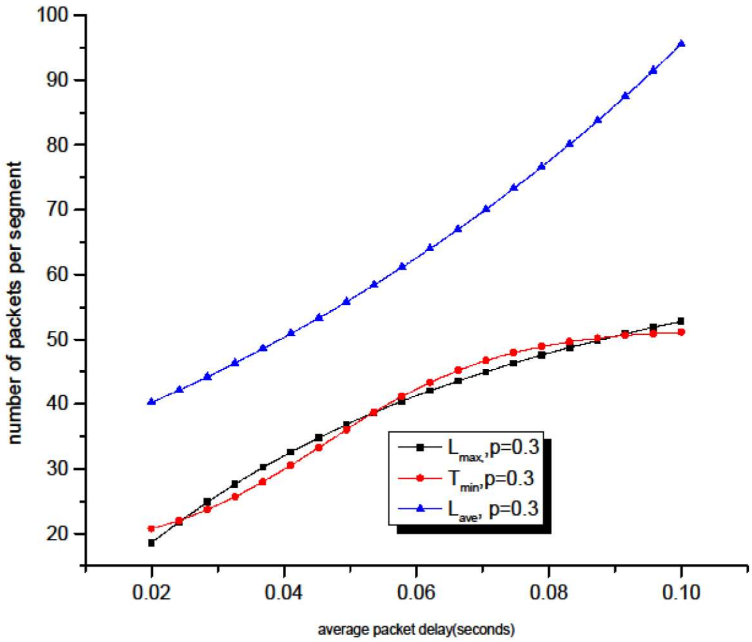

Figure 5-27: Segment size versus average packed delays

In carrying out the simulation traffic sources generated identical connection requests. We observe from the graph that for a fixed average delay, the T_{ave} scheme assembles more packets per segment, thus, resulting in longer bursts, as well as relatively lesser processing overheads at the nodes because of the lesser corresponding control packets arriving rates. HP and LP segment sizes versus arrival is plotted in figure 5-28. As can be observed at $\rho = 0.7$, the average segment length is about 40, which we will now assume throughout the rest of this evaluation. In the core nodes, the composite burst scheduling as well as RWA is accomplished according to the algorithm summarised next

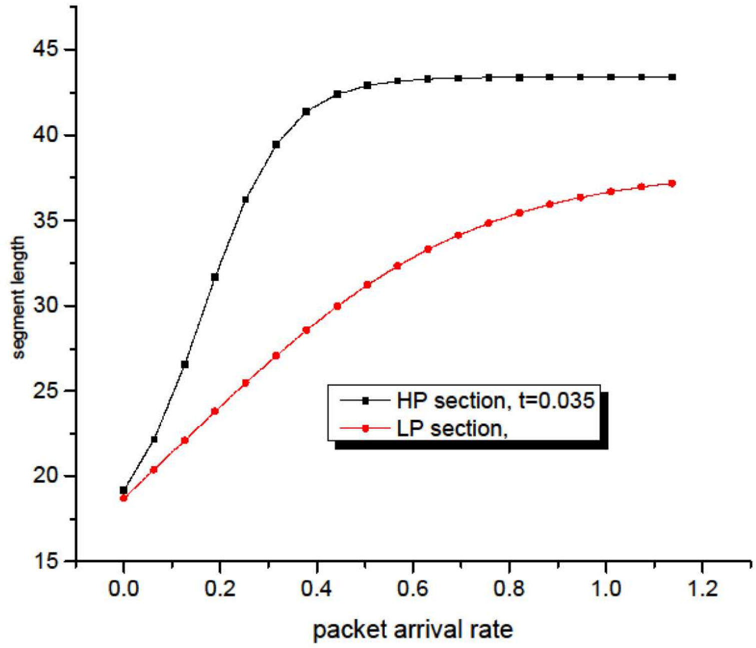

Figure 5-28: Segment size comparisons for both HP and LP segments versus packet arrival rate

The following is a pseudo code of the implementation of the proposed approach.

1. % **initialise I:** arrival of a CHP.

2 acquire: $\lambda_{in}^{l} \leftarrow$ incoming composite burst's wavelength;

3 $P_{in} \leftarrow$ incoming port of composite burst,

4 L, composite burst length;

5 L_{lp}-LP section size, Γ_{o}-composite burst arrival time;

6 % **contention check**

7 **if** contention(λ_{in}^{l} _ P_{in} _, L, Γ_{o}) **then**

8 Can reschedule contending section $\leftarrow$ *True*

9 //Priority $\leftarrow$ *HP* contending section//;

10 if Can reschedule then

11 reserve new_ λ _for truncated section, go to 13

12 else

% **:random alternate route calculation**

13 **Initialize,** $\rho_{l,j} = 0, \forall$ nodes j .

14 **for all** i **, do &&** for all paths, $p \in P(i)$; $\Phi_{p,i} = \lambda_{p,i}$

15 P :=random solution

16 Evaluate F of P

17 *return* (F, V)

end.

HP packets constitute 25% of all arrivals.

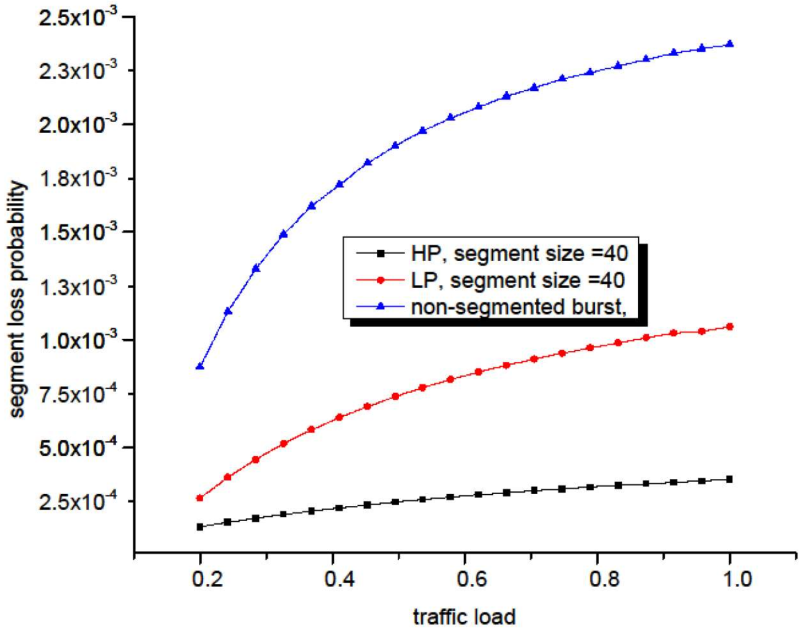

Figure 5-29: Burst/Segment Loss probabilities

Figure 5-29, plots the burst/segment loss probability for both HP and LP segments. This is compared with the conventional length based OBS approach in which all traffic type have the same priority. The contending composite burst section will pre-empt the original burst on the original route (first choice route) if its priority is higher. By fixing the traffic load to 0.6, we observe from the same graph that the HP segments have a lower blocking probability ($< 2.5\times10^{-4}$), partly because of their default priority. We also observe that in general the segment loss probability is generally lower for the HP/LP composite bursts since the overall burst size is fixed, whereas non-segmented burst approach has exponentially distributed burst sizes. The average end-to end delay is plotted against traffic load (ρ) in Figure 5-30. Three schemes cases are compared namely; (i) when contention resolution is by way of using FDLs coupled with deflection routing; (ii) no deflection and (ii the proposed approach. It is generally observed that the average end-to-end delay peaks at about $\rho = 4$ and then afterwards it starts to decrease. This is partly because of lesser SLE effects at higher network loads and that bursts traversing longer hops are likely to be discarded along the way (prior to reaching the destination) as the input load increases.

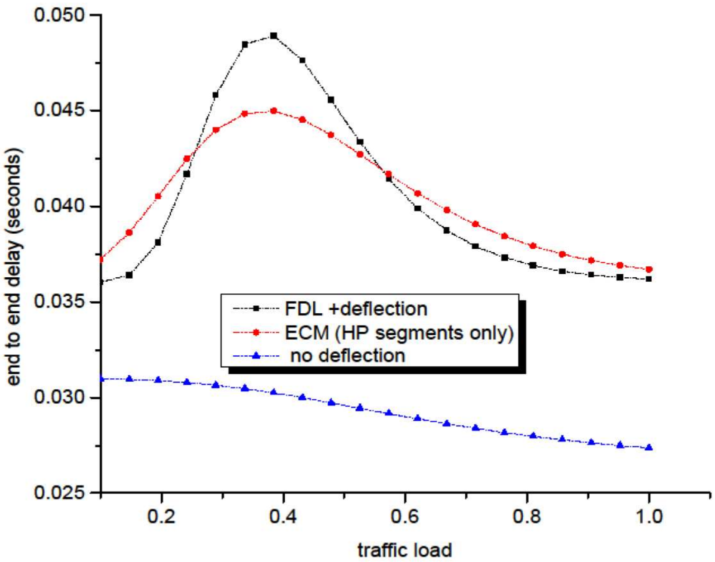

Figure 5-30: Average end to end delays

Note that at 45% (0.45) network traffic loading, the ECM scheme lowers delays to about 0.045 seconds, which is less that of the existing FDL+ deflection schemes (0.05 seconds). The case of no deflection is impractical as there will be excessive congestion levels.

Also it is noted that for low to medium network traffic loads, deflection routing increases the overall number of hops. It is also observed that the average end-to-end delay of the proposed approach decreases at low network loads.

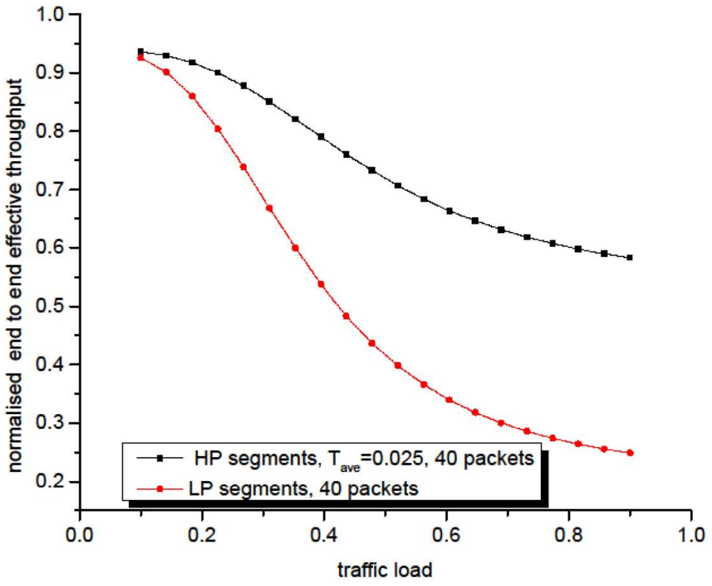

Figure 5-31: Normalized end to end effective throughput

We further evaluate the extent to which our proposed approach, will promote efficient use of available network resources. In this regard, we first define "effective throughput" as the average number of composite bursts traversing the end to end path S to d per unit time. In Figure 5-31 we plot the normalised average effective throughput as a function of the network traffic load. In general, the average effective throughput in the proposed approach gradually increases with increase in network traffic, whereas the same decreases in the conventional OBS with deflection routing.

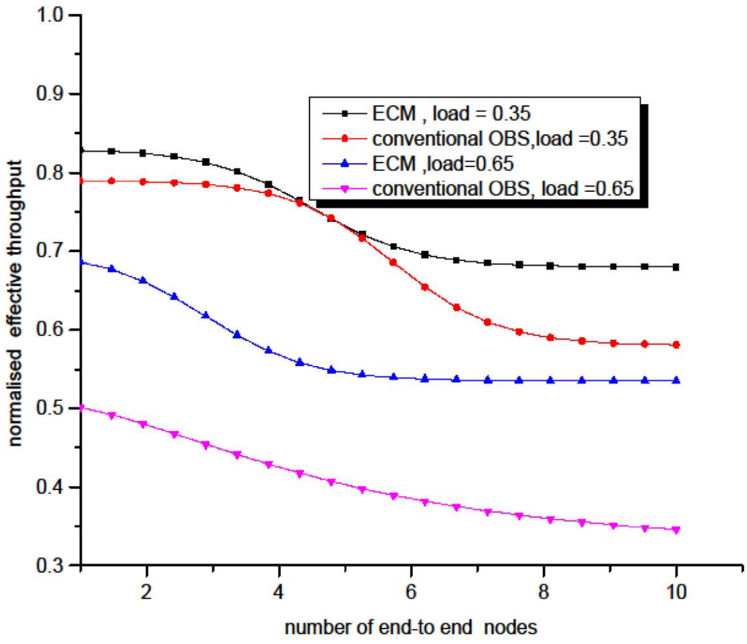

Figure 5-32: Normalized effective throughput.

In Figure 5-32, when the load is set to 0.35, and the number of traversed nodes is set to 3,, the ECM scheme yields an effective throughput of 0.84, whereas conventional OBS's effective throughput is less than 0.8. . When the number of nodes is increase to 9, still the ECM outperforms (0.7). When the network load is increase to 0.65, we see that the ECM's effective throughput is comparably higher. Conclusively as can be seen from the Figure 5-32, our approach still enables the network to utilise resources efficiently when the network is congested. The same trend depicts itself when the average effective throughput is plotted as a function of the number of hops on the first choice route.

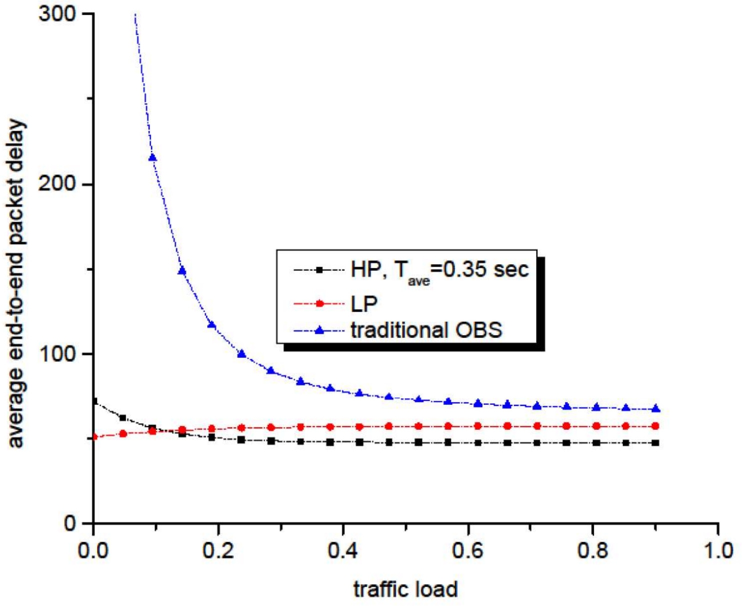

Figure 5-33: Overall end-to-end latency versus network load-traffic load.

Finally, we evaluated the overall end to end latency for HP ($T_{ave} = 0.35 sec$) versus LP traffic. We observe that both traffic classes incur acceptable delays at very low traffic loads and that at moderate to high network loads, HP traffic incurs the least delays (latencies)

5.4.4 Conclusion

In this work, we have proposed and evaluated a congestion management approach we refer to as "enhanced congestion management" which gears towards rendering and guaranteeing a consistent QoS as well as rational and fair use of available links. This is primarily achieved by way of combining priority service differentiation, time averaged delay segmented burstification load based random shortest deflection path selection. The proposed approach does not add any extra hardware complexity to existing OBS network interior nodes, but provisions a variable offset time t_i which will facilitate priority based contention resolution in the interior network hence minimising its effects as well as congestion. In that way, the available network resources are appropriated more fairly and efficiently. In comparison to similar conventional OBS approaches, the proposed approach improves the network performance in terms of loss probability when the network is not so highly loaded. Simulation results confirm these properties and also show the effectiveness of the proposed method with a higher intensity of high priority traffic. The good put (criteria of network capability of transporting useful packets) is increased due to the recovery of lost bursts.. Future works are finding the numerical model and applying this result to inter-class cases. Also, the adoption of combined routing strategies depending on traffic load variations can result in better performance. Another improvement can exploit the combined use of this static routing approach with the three basic deflection domains (time, space and wavelength) for contention resolution.

Chapter 6. Conclusion and Future Research

6.1 Summary

Current global data traffic is increasingly dominated by delay and loss intolerant IP traffic which generally displays a structural self-similarity. Due to the buffer-less nature of optical burst switched (OBS) networks, contention/congestion in the core network can quickly lead to degradation in overall network performance at moderate to high traffic levels due to heavy burst loses. Several approaches have been explored to address this problem, notably measures that would minimize burstification delays, congestion, blocking at the same time enhancing end-to-end throughput as well as rational and fair utilization of the links. In this way contention, congestion and consequently QoS degradation for delay as well as loss sensitive applications data as it traverses such networks can be minimized. Network implementation of carefully designed routing and wavelength assignment algorithms will drastically reduce burst losses, enhance links utilization and promote acceptable fairness to all traffic types. This is likely to promote better chances of a guaranteed consistent QoS. Several studies have focused on end to end blocking probability in such networks as well as utilization. Utilization of the links is studied in detail in which the concepts of effective utilization and ineffective utilization are used as key performance measures of OBS networks. Effective utilization relates to channels used by bursts that eventually reach their destinations, whereas ineffective utilization relates to channels that are used by bursts that are discarded before reaching their ultimate destinations. Reduction of the latter has the effect of improving both the performance and efficiency of the network. Ineffective utilization is particularly more pronounced in periods of heavy traffic loads. Under heavy traffic load conditions, the effective utilization and network throughput consequently suffer congestion collapse, as in this case, the total network utilization tends to be quite high, even though ineffective utilization dominates, i.e. most of the utilized capacity is used by bursts that are eventually discarded before reaching the intended destinations.

Noting that congestion and contention minimization are key to a consistent QoS provisioning the work herein focused on addressing the two. In particular, we focused on analyzing mechanisms and approaches geared towards alleviating congestion and con to review OBS switching paradigm, partially in motivation it as the ultimate backbone network switching for NGNs. As such the work addressed the following key issues

- Burstification schemes taking into account the diverse QoS requirements of various applications.
- Existing contention and congestion resolution approaches, as well as various signaling scenarios.
- Routing and wavelength assignment/allocation approaches in OBS networks.
- An evaluation of a congestion and contention minimization scheme that aims improving QoS of real time applications.

The scheme we proposed and evaluated is primarily service differentiation based and overall aimed at congestion, blocking and latency minimization, by way of combining time averaged delay segmented burstification as well as random shortest path selection based deflection routing and wavelength assignment. Simulation results showed that the scheme can effectively minimize congestion and at the same time improve both throughput and effective utilization, under moderate to high network traffic conditions. Overall, we show that the approach guarantees a consistent QoS.

In our future work, we will focus on novel methods that considerably enhance throughputs at the same time reducing burst losses using traffic prediction models. Notably from a viewpoint of flow control, we would postulate a burst flow control approach in which regulates the length of a burst based on the congestion weight on its path. In that way, can achieve improved efficiency and fairness compared to the case where a burst contention resolution scheme is solely used. Developing an edge dynamic delay method which decreases burst losses especially when two or more bursts are imminently contending in the core nodes will also be explored.

Published Papers

[1]. Ndadzibaya Gomba, Bakhe Nleya, Andrew Mutsvangwa. A Congestion Minimization Strategy For OBS Backbone Networks. Proceedings of the IEEE's International Conference, September 14-17, Addis Ababa.

[2]. Ndadzibaya Gomba, Bakhe Nleya, Andrew Mutsvangwa. Class Segmented Burst Assembling Approach in OBS Networks. Proceedings of the IEEE's International Conference, September 14-17, Addis Ababa.

[3]. Ndadzibaya Gomba, Bakhe Nleya, Andrew Mutsvangwa. A Congestion Avoidance/Minimization Approach for Optical Burst Switched Backbone Networks. Proceedings of Southern Africa Telecommunication Networks and Applications Conference (SATNAC) 2015.

[4]. Ndadzibaya Gomba, Bakhe Nleya, Priority Burst Segmentation for Congestion Management in OBS Networks. IEEE"s International Conference on Advances in Computing, Communication and Engineering, 28-29 November, 2016

.

References

[1] http://www.itu.int/en/ITU-T/gsi/ngn/Pages/definition.aspx..

[2] Jongtae Song; Mi Young Chang; Soon Seok Lee; Jinoo Joung; " Overview of ITU-T NGN QoS Control", Communications Magazine, IEEE , September 2007,Vol.45 Issue 9, pp 116 – 123.

[3] Huang Yiding; Li Jian; Wan Shuqin; "Technology and prospects of NGN", Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on, April 2010, pp 90 – 92, ISBN: 978-1-4244-5263-7.

[4] ZhiGang Sun, Yi Dai and ZhengHu Gong "MPFS: A truly scalable router architecture for next generation Internet", Science in China Series F: Information Sciences, Vol 51, N0. 11, 1761-1771, DOI: 10.1007/s11432-008-0148-4

[5] Y. Chen, C. Qiao, and X. Yu, "Optical Burst Switching (OBS): A New Area in Optical Networking Research", Network, IEEE, 2004 – ieeexplore.org

[6] Y. Xiong, M. Vandenhoute, and H. C. Cankaya, "Control architecture in optical burst-witched WDM networks," IEEE J. Sel. Areas Commun., Vol. 18, no. 10, pp. 1838–1851, Oct. 2000.

[7] Skorin-Kapov, Nina, "Multicast routing and wavelength assignment in WDM networks: a bin packing approach", *Journal Of Optical Networking* Volume: 5 Issue: 4 ISSN: 1536-5379

[8] H. Zang, J. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical Wavelength Division Multiplexing networks", *Optical Networks Magazine*, 1(1):pp. 47-60, January 2000.

[9] Z. Rosberg, H. L. Vu, M. Zukerman, and J. "White, Blocking probabilities of optical burst switching networks based on reduced load fixed point approximations", In *Proceedings of IEEE INFOCOM 2003*, New York, NY (USA), March-April 2003.

[10] Gee-Kung Chang, Jianjun Yu, Yong-Kee Yeo, A. Chowdhury, J. Zhensheng, "Enabling technologies for next-generation optical packet-switching networks," Proceedings of the IEEE, Vol. 94, No. 5, pp.892 - 910, May 2006.

[11] J. Yu, G. K. Chang, A. Chowdhury and J. L. Long, "Spectral efficient DWDM optical label/payload generation and transport for next generation internet," IEEE Journal of Light wave Technology, 2004, Vol. 22, no. 11, pp. 2469-2482.

[12] Chowdhury, J. Yu and G. K. Chang, "A Novel Optical Label Swapping Scheme for DPSK Data Transmissions Using Optical Carrier Suppression and Separation Technique without requiring a conventional wavelength converter," Optical Fiber Communication Conf. (OFC) 2005, OTuC4.

[13] C. Gauger, "Contention resolution in Optical Burst Switching networks," in Advanced Infrastructures for Photonic *Networks: WG 2 Intermediate Report, 2002, pp. 62–82.*

[14] F. Masetti et.al. "Design and Implementation of a Multi-Terabit Optical Burst/Packet Router prototype," in OFC *Post deadline Papers, 2002, pp. FD11–FD13.*

[15] J. Ramamirtham and J. Turner, "Design of Wavelength Converting Switches for Optical Burst Switching," in *Proceedings of INFOCOMM, 2002,* Vol. 1, pp. 362–370

[16] R. S. Tucker and W. De Zhong. "Photonic packet switching: An overview". *IEICE Trans. commun*, E82-B (2): pp. 254-264, February1999.

[17] J. Cheyns, C. Develder, E. Van Breusegem, A. Ackaert, M. Pickavet, and P. Demeester. "Routing in an awg-based optical packet switch". *Photonic Network Communications*, 5(1): pp. 69-80, 2003.

[18] T. Coutelen, H. Elbiaze, and B. Jaumard. "Performance comparison of ocs and obs switching paradigms". In *Proceedings of the 7th IEEE International Conference on Transparent Optical Networks (ICTON2005)*, Barcelona, Spain, July 2005.

[19] Agrawal, T. S. El-Bawab, and L. B. Sofman. "Comparative account of bandwidth effeciency in optical burst switching and optical circuit switching networks". *Photonic Network Communications*, 9(3): pp. 297-309, 2005.

[20] Xin Liu, Chunming Qiao, Xiang Yu, and Weibo Gong. "A fair packet-level performance comparison of OBS and OCS". In *Proceedings of Optical Fiber Communication (OFC 2006)*, Anaheim, CL (USA), March 2006.

[21] Angeliki Sideri and Emmanouel A. Varvarigos, "New Assembly Techniques for Optical Burst Switched Networks Based on Traffic Prediction" Computer Engineering and Informatics Dept, University of Patras, Greece

[22] Kantarci, Burak,Oktug, Sema F.,Atmaca, Tülin. "Performance of OBS techniques under self-similar traffic based on various burst assembly techniques" ,SN - 0140-3664 doi: DOI: 0.1016 j.comcom.2006.08.035

[23] Amit Kumar Garg, "Analysis of burst/packet assembly techniques in high-speed optical switching network", Optik 122 (2011) 616–619, journal homepage: www.elsevier.de/ijleo

[24] K. Christodoulopoulos, E. Varvarigos and K. Vlachos, "A new Burst Assembly Scheme based on the Average Packet Delay and its Performance for TCP Traffic", under 2nd revision, Optical Switching and Networking, Elsevier, 2006.

[25] M. Yoo and C. Qiao, "Just-Enough-Time (JET): A high speed protocol for bursty traffic in optical networks," in *IEEE/LEOS Technologies for a Global Information Infrastructure*, Aug. 1997, pp.26–27.

[26] J. Wei, J. Pastor, R. Ramamurthy, and Y. Tsai, "Just-in-time optical burst switching for multiwavelength networks," in *IFIP Broadband Commun.*, Hong Kong, Nov. 1999, pp. 339–352

[27] J. Wei and R. McFarland, "Just-in-time signalling for WDM optical burst switching networks," *Lightwave Technol.*, Vol. 18, no. 12, pp. 2019–2037, Dec. 2000.

[28] P. Bayvel, "Wavelength routing and optical burst switching in the design of future optical network architectures," in *European Conf. on Optical Commun.*, Vol. 4, 2001, pp. 616–619.

[29] Yuhua Chen Turner, J.S. Pu-Fan Mo , "Optimal Burst Scheduling in Optical Burst Switched Networks" , journal of lightwave technology, Vol. 25, no. 8, august 2007.

[30] J. Xu, C. Qiao, J. Li, and G. Xu, "Efficient channel scheduling in optical burst switched networks," in Proc. IEEE INFOCOM, 2003, Vol. 3, pp. 2268–2278.

[31] Kaheel, T. Khattab, A. Mohamed, and H. Alnuweiri." Quality of Service Mechanisms in IP-over WD Network". IEEE Communications Magazine. December, 2002.

[32] V.M. Vokkarane and J. P. Jue, "Segmentation-Based Non-Preemptive Scheduling Algorithms for Optical Burst-Switched Networks", Proc. First International Workshop. OBS, Oct. 2003.

[33] tp://www.cisco.com/en/US/docs/net_mgmt/ip_solution_center/3.0/qos/user/guide/concepts. Html Last accessed on 30th January 2014 – 1509 hrs.

[34] R.Sivaraman, V.R.Sarma Dhulipala, L.Sowbhagya, B.Vishnu Prabha, "Comparative Analysis of QoS Metrics in Mobile Ad Hoc Network Environment", International Journal of Recent Trends in Engineering, Vol. 2, No. 4, November 2009

[35] N .Akar, E .Karasan, K. G .Vlachos, E. A Varvarigos, D. Careglio, M. Klinkowski, J. Sole-Pareta . "A survey of quality of service differentiation mechanisms for optical burst switching networks" ,Optical Switching and Networking ISSN: 1573-4277 Vol: 7 (1) 2010 pp. 1-11

[36] I. De Miguel, J. C. Gonzalez, T. Koonen, R. Duran, P. Fernandez, and I. T. Monroy. "Polymorphic architectures for optical networks and their seamless evolution towards next generation networks." Photonic Network Communications, 8(2), pp.177-189, 2004.

[37] K. Dolzer and C.M. Gauger. "On burst assembly in optical burst switching networks - a performance evaluation of just-enough-time". In Proceedings of the 17th International Teletraffic Congress (ITC 17), Salvador, Brazil, December 2001.

[38] N. Barakat, E. Sargent, "On optimal ingress treatment of delay sensitive traffic in multi-class OBS systems", in: Proceedings of 3rd International Workshop on Optical Burst Switching, WOBS 2004, collocated with Broad Nets, San Jose, CA, USA, 2004.

[39] R. Froom, M. Flannagan, K. Turek, "Cisco Catalyst QoS: Quality of service in Campus Networks", Cisco Press, June 2003.

[40] S. Floyd, V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking, Vol.1, p. 397 – 413, 1993.

[41] J. Gettys "Red in a Different Light", jg's Ramblings, May 2012.

[42] V. Dadarlat, R. Jeler, A. Peculea, B. Iancu, E. Cebuc, C. Ardelean. "Development Framework for Congestion Avoidance Mechanisms", Acta Technica Napocensis (ProQuest Indexed), ISSN 1221-6542, Vol. 51, No. 3, 2010.

[43] A. Varga, OpenSim Ltd., "OMNeT++ User Manual", http://www.omnetpp.org, 2011.

[44] S. Bajaj, L. Breslau, D. Estrin, K. Fall, S. Floyd, M. Haldar, P. Handley, A. Helmy, J. Heidemann, P. Huang, S. Kumar, S. McCanne, R. Rejajie, S. Punnet, K. Varadhan, X. Ya, H. Yu, D. Zapalla "Improving simulation for network research" Technical Report 99-703b, USC Computer Science Department, 1999.

[45] OPNET Technologies, Inc. OPNET Modeler, http://www.opnet.com

[46] J-Sim home page, Available: http://www.j-sim.org.

[47] A. Varga, R. Hornig, "An Overview of the OMNeT++ Simulation Environment", In 1st ACM/ICST International Conference on Simulation Tools and Techniques for Communications, Networks and Systems (SIMU Tools 2008), Marseille, France, March 2008.

[48] V. Jacobson, M. Karels, "Congestion Avoidance and Control", Proceeding of ACM SIGCOMM'88 Symposium, 1988.

[49] S. Floyd, R. Gummadi, S. Shenker "Adaptive RED: An Algorithm for Increasing the Robustness of RED's Active Queue Management",

[50] "INET Framework", http://inet.omnetpp.org/ , March 2012.

[51] Z. Wang, "Internet QoS: Architectures and Mechanisms for Quality of Service", Morgan Kaufmann, San Francisco, p. 64, 2001.

[52] K. Christodoulopoulos, E. Varvarigos and K. Vlachos, "A New Burst Assembly Scheme Based on the Average Packet Delay and its Performance for TCP Traffic, "Elsevier Optical Switching and Networking, Vol. 4, iss. 3-4, pp.200-212, November 2007.

[53] A. Sideri and E. A. Varvarigos, "New Assembly Techniques for Optical Burst Switched Networks Based on Traffic Prediction," Lecture Notes in Computer Science, Proc. ONDM, Vol. 4534, pp. 358-367, 2007.

[54] B. Kantarci and S. Oktug, "Adaptive Threshold Based Burst Assembly in OBS Networks," Proc. Canadian Conference on Electrical and Computer Engineering (CCECE 2006), pp. 485-488, Ottawa, Canada, May 2006.

[55] B. Kantarci and S. Oktug, "Path Loss Rate Driven Burst Assembly in OBS Networks," Lecture Notes in Computer Science, Proc. ISCIS, Vol. 4263, pp. 483-492, 2006.

[56] J. Zhou, J. Wu and J. Lin, "Improvement of TCP Performance over Optical Burst Switching Networks," Lecture Notes in Computer Science, Proc. ONDM, Vol. 4534, pp. 194-200, July 2007.

[57] S. Peng, Z. Li, X. Wu and A. Xu, "TCP Window Based Dynamic Assembly Period in Optical Burst Switching Network," Proc. IEEE ICC'07, pp. 2365-2370, June 2007.

[58] J. F. Kurose and K. W. Ross, "Computer Networking: A Top-Down Approach Featuring the Internet." Addison-Wesley, 2005.

[59] J. Wu and H. El-Ocla, "TCP Congestion Avoidance Model with Congestive Loss," Proc. IEEE ICON'04, Vol. 1, pp. 3-8, November 2004.

[60] C. Jin, D. X. Wei, S. H. Low, G. Buhrmaster, J. Bunn, D. H. Choe, R. L. A. Cottrell, J. C. Doyle, W. Feng, O. Martin, H. Newman, F. Paganini, S. Ravot, and S. Singh. “FAST TCP: From theory to experiments”. IEEE Network, 19(1) pp. 4–11, January/February 2005.

[61] T. Kelly. “Scalable TCP: Improving performance in high speed wide area networks”. In Proceedings of PFLDnet, Geneva, Switzerland, Feb. 2003. [11] D. Leigh, R. Shorten, et al. H-TCP ns-2 implementation, 2005.

[62] The Network Simulator. http://www.isi.edu/nsnam/ns/.

[63] I. Rhee and L. Xu. “CUBIC: A new TCP-friendly high speed TCP variant”. In Proceedings of PFLDnet, Lyon, France, Feb. 2005.

[64] I. Rhee and L. Xu. “Simulation code and scripts for CUBIC”, 2005.

[65] C. Jin, D. Wei, S. H. Low, G. Buhrmaster, J. Bunn, D. H. Choe, R. L. A. Cottrell, J. C. Doyle, H. Newman, F. Paganini, S. Ravot, and S. Singh, “FAST Kernel: Background theory and experimental results,” in First International Workshop on Protocols for Fast Long-Distance Networks, February 2003..

[66] E. Souza and D. A. Agarwal. “A High Speed TCP study: Characteristics and deployment issues”. Technical Report LBNL-53215, 2003.

[67] K. Tokuda, G. Hasegawa, and M. Murata. “Performance analysis of High Speed TCP and its improvements for high throughput and fairness against TCP Reno connections”. In High-Speed Networking Workshop (HSN), 2003.

[68] L. Xu, K. Harfoush, and I. Rhee. “Binary increase congestion control for fast, long distance networks”. In Proceedings of IEEE INFOCOM, Hong Kong, Mar. 2004.

[69] S. Floyd. “High speed TCP for large congestion window”, Internet Draft draft-floyd-tcp-highspeed-01.txt, February 2003.

[70] S. Floyd and K. Fall. “Promoting the use of end-to-end congestion control in the Internet”. IEEE/ACM Transactions on Networking, 7(4):458–472, August 1999.

[71] Eric He, Michael Welzl and Pascale Vicat-Blanc Primet “A Survey of Transport Protocols other than Standard TCP” Data Transport Research Group, GFD-I.055, 2005.

[72] http ://www.cs.ucla.edu/NRL/hpi/tcpw/tcpw_papers/part1-congestion-control.pdf.

[73] Eric He, Rajkumar Kettimuthu, Sanjay Hegde, Michael Welzl, Jason Leigh, Chaoyue Xiong and Pascale Vicat-Blanc Primet, “Survey of Protocols and Mechanisms for Enhanced Transport over Long Fat Pipes”, Data Transport Research Group,2003/2004.

[74] D. M. Lopez-Pacheco and C. Pham "Robust Transport Protocol for Dynamic High-Speed Networks: enhancing the XCP approach" in Proc. of IEEE MICCICON, 2005.

[75] Sumitha Bhandarkar, Saurabh Jain and A. L. Narasimha Reddy, "Improving TCP Performance in High Bandwidth High RTT Links Using Layered Congestion Control", International Workshop on Protocols for Fast Long-Distance Networks, February 2005.

[76] Sangtae Ha, Long Le, Injong Rhee and Lisong Xu "Impact of background traffic on performance of high-speed TCP variant protocols" S. Ha et al., Computer Networks 51, pp. 1748–1762, 2007.

[77] Sangtae Ha, Yusung Kim, Long Le, Injong Rhee, and Lisong Xu* "A step toward realistic evaluation of highspeed TCP protocols", 2006.

[78] Yunhong Gu, Xinwei Hong, and Robert L. Grossman, "Experiences in Design and Implementation of a High Performance Transport Protocol", Supercomputing, Proceedings of the ACM/IEEE SC2004 Conference, pp. 22, 2004.

[79] D. J. Leith and R. Shorten. "H-TCP Protocol for High-Speed Long Distance Networks".PFLDnet '04, Feb. 2004.

[80] Pankaj Sharma, "Performance Analysis of High-Speed Transport Control Protocols", Master of Science (Computer Science) Thesis, Clemson University, August 2006.

[81] Michele C. Weigle, Pankaj Sharma, and Jesse R. Freeman IV, "Performance of Competing High-Speed TCP Flows" Networking 2006. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems, pp. 476-487, 2006.

[82] Yee-Ting Li, Douglas Leith and Robert N. Shorten, "Experimental Evaluation of TCP Protocols for High-Speed Networks", IEEE/ACM Transactions on Networking, pp.: 1109 – 1122, Oct. 2007.

[83] Sally Floyd, Vern Paxson, "Difficulties in simulating the Internet", IEEE/ACM Transactions on Networking 9 (4), 2001.

[84] Park, WS, Shin, M, Lee, HW, and Chong, S. A Joint Design of Congestion Control and Burst contention Resolution for Optical Burst Switching Network. Light wave Technology, Vol. 21, Number 17, 2009.

[85] H. Zang, J. P. Jue, and B. Mukherjee, "A Review of Routing and Wavelength Assignment Approaches for Wavelength-Routed Optical WDM Networks," Optical Networks Magazine, Vol. l, Number 1, pp. 47-60, 2000.

[86] Shakeel Ahmad, Adli Mustafa1, Bashir Ahmad, Arjamand Bano and Al-Sammarraie Hosam. Comparative Study of Congestion Control Techniques in High Speed Networks. International Journal of Computer Science and Information Security, Vol. 6, Number. 2, 2009.

[87] Li Yang and George N. Rouskas. Adaptive Path Selection in OBS Networks. Journal of Lightwave Technology, Vol. 24, Number 8, pp3002-, 2006.

[88] Vardi, Y. Network Tomography: Estimating Source-Destination Traffic Intensities from Link Data Journal of the American Statistical Association March 1, 1996.

[89] Shuo Li , Moshe ,Zukerman, Meiqian Wang, Eric W.M. Wong. Improving Throughput and Effective Utilization in OBS networks. Optical Switching and Networking. Vol. 18, Part 3, November 2015, Pages 222–234.

[90] Meiqian Wang, Shuo Li, Eric W. M. Wong and Moshe Zukerman. Evaluating OBS by Effective Utilization. IEEE Communications letetrs, Vol. 17, Number. 3, March 2013.

[91] Zvi Rosberg_, Hai Le Vu†, Moshe Zukerman† and Jolyon White. Blocking Probabilities of Optical Burst Switching Networks Based on Reduced Load Fixed Point Approximations. IEEE INFOCOM , March 30-April3, 2003.

[92] V. Kavitha andV. Palanisamy. New Burst Assembly and Scheduling Technique for Optical Burst Switching Networks. Journal of Computer Science, Vol. 9, Number 8, pp 1030-1040, 2013.

[93] Xinhui Xu, Chunming Qiao, Jikai Li, and Guang Xu. Efficient Channel Scheduling Algorithms in Optical Burst Switched Networks. IEEExplore, 2003.

[94] M. Kozaka, B. Jaumard, Boha. On the Efficiency of Stream Line Effect for Contention Avoidance in Optical Burst Switching Networks. Optical Switching and Networking Vol.18, Part 1, November, Pages 35–50, 2015.

[95] Amit Kumar Garg .Congestion Prevention for Reliable OBS. Optik - International Journal for Light and Electron Optics. Volume 124, Issue 21, November, Pages 5025–5029, 2013.

[96] URI Zwick. All Pairs Shortest Paths Using Bridging Sets and Rectangular Matrix Multiplication. Zabinsky Z. B.Journal of the ACM, Vol. 49, Number 3, May, pp 289-317,2002. .

[97] Ulrich Meyer. Design and Analysis of Sequential and Parallel Single–Source Shortest–Paths Algorithms. Universit¨at des Saarlandes, Saarbr¨ucken, 2002.

[98] E. Altman, K. Avrachenkov, C. Barakat, A.A. Kherani, B.J. Prabhu "Analysis of MIMD congestion control algorithm for High Speed Networks" Computer Networks: The International Journal of Computer and Telecommunications Networking Vol. 48 , Issue 6, pp.: 972 – 989, 2005.

[99] For XCP: http://www.isi.edu/nsnam/ns/doc/node238.html

[100] R. El Khoury and E. Altman "Analysis of Scalable TCP" In Proceedings of the International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks (HET-NETS), July 2004. http://citeseer.ist.psu.edu/719259.html