A Bursts Contention Avoidance Scheme Based on Streamline Effect Awareness and Limited Intermediate Node Buffering in the Core Network

Bakhe Nleya and Richard Chidzonga
Department of Electronic Engineering
Durban, South Africa
bmnleya@gmail.com/ rfchidzonga@gmail.com

Abstract— **In an Optical Burst Switched (OBS) network, data packets sourced from peripheral networks are assembled into huge sized data bursts. For each assembled data burst, an associated control signal in the form of a burst control packet is (BCP) is generated and scheduled at an offset time ahead of the data burst. The offset timing is to allow for the pre-configuration of required resources at all subsequent intermediate nodes prior to the actual data burst's arrival. In that way, the data burst will fly-by each node and hence no requirement for temporary buffering at all intermediate nodes. An operational requirement of an OBS network is that it be loss-less as in that way a consistent as well as acceptable quality of service (QoS) for all applications and services it serves as a platform can be guaranteed. Losses in such a network are mainly caused by improper provisioning as well as dimensioning of resources thus leading to contentions among bursts and consequently discarding of some of the contending data bursts. Key to both provisioning as well as proper dimensioning of the available resources in an optimized way is the implementation of effective routing and wavelength (RWA) that will seclude any data losses due to contention occurrences. On the basis of the effects of the streamline effect (SLE), that is, effectively secluding primary contention among flows (streams) in the network, we propose in this paper a limited intermediate buffering that couples with SLE aware prioritized RWA (LIB-PRWA) scheme that combats secondary contention as well. The scheme makes routing decisions such as selection of primary and deflection routes based on current resources states in the candidate paths. A performance comparison of the proposed scheme is carried out and simulation results demonstrate its comparative abilities to effectively reduce losses as well as maintaining both high network resources utilization as well as QoS.**

Keywords — Optical burst switching, streamline effect, congestion, contention, routing and wavelength assignment.

I. INTRODUCTION

The emergency of Internet of Things (IoT) enabled networks has resulted in a surge of various applications and services and generating massive amounts of traffic globally. This is necessitating the design and deployment of an all optical transport network infrastructure to serve as the core backbone network for the resultant diverse communications services. Such an infrastructure provides connectivity to millions of administrative, commercial, industrial as well as residential centers. The heterogeneous nature of the large volumes of traffic generated by various applications and services ideally requires an all-optical backbone network infrastructure to accommodate it. Such a network must be continuously adaptable to the changing nature of the traffic as well as its spontaneous growth with time. In so doing, it must ensure high end-to-end QoS, availability as well as provision adaptable controllability in cooperation with peripheral (service) layer networks. Utilization of dense wavelength division multiplexing (DWDM) in optical fibers has resulted in transmission bearers achieving speeds in the order of Terabits per second. However, current router switches have not been able to solve the speed mismatches between the high DWDM transmission speeds versus their low switching capabilities. Optical burst switching (OBS) is being rolled out to narrow the switching versus transmission speed gaps in current and future generation optical backbone networks. The OBS approach is based on aggregating and assembling data packets at ingress nodes into optical data bursts. A control packet (CP) is separately generated to carry control data for each assembled burst on a separate wavelength channel. It is transmitted ahead of the actual burst and will thus reach the next intermediate node within some preset offset-time [1], [2].

The magnitude of this timing is carefully set so that it is sufficient to allow for the CP's processing by a CP controller at all intermediate nodes. This also allows for the node's switch fabric pre-configuring as well as channel reservation on its output link prior to the actual arrival of the optical data burst. This prior reservation of resources eradicates the need for optical burst buffering during the switching process, otherwise this would escalate network design and operational costs. The optical burst is then switched through and its reserved resources freed and made available for other lightpath connection requests. The OBS switching paradigm is prone to both congestion and contentions. Both reactive and proactive measures may be employed in the network to try to avoid contention. Such measures include backpressure routing, network segmentation, as well as prioritizing the network traffic. However, the existence of any congested links may drastically aggravate network throughput and consequently its overall performance. Notable QoS metrics that degrade as a consequence of congestion occurrences are burst blocking rate and end-to-end latency.

Burst contentions occurring in the core nodes may lead to some data bursts being deleted as a resolution measure. Overall, given the limited buffering at the core nodes, it is necessary that contention and congestion avoidance be jointly implemented in order to improve network throughput, thus in the process guaranteeing consistent acceptable QoS for the various applications and services. Burst assembling approaches at ingress nodes, RWA, contention/congestion resolution are key to minimizing both contention as well as congestion.

II. AN OBS NODE

In order to transmit data bursts in an OBS network, lightpath connections are setup between desired source and destination pairs. A typical lightpath connection request is established through a series of lightpath connections from source to destination. These will now accommodate both control data as well as the data bursts. At each optical node, functionalities such as multiplexing and demultiplexing of channels as well as wavelength routing should be supported.

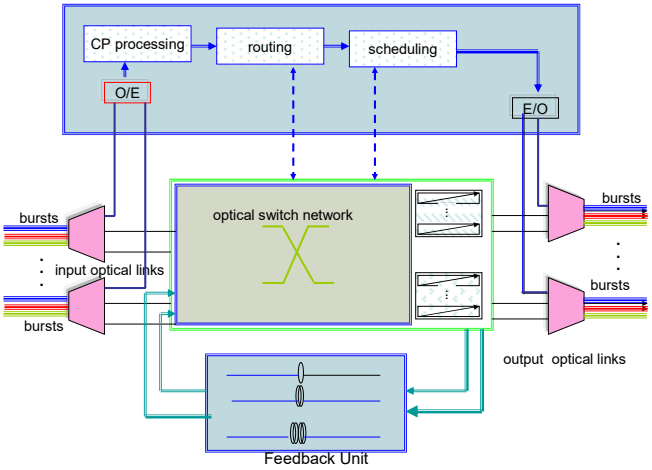

Figure 1. An Example OBS core node

As the data bursts are switched to intended output ports at network nodes, contentions may occur. It is therefore necessary to provision contention resolution mechanisms that will ensure burst loss minimization. We therefore assume that the architecture of each node should be designed in conformity with the operations of the priority-based intermediate buffering and routing and wavelength assignment (RWA) scheme (LIB-PRWA) which we shall describe in due course.

The logical architecture of such a node is shown in Figure 1. Typically, the edge-core joint node example is a composite edge and core nodes. Such an architecture can perform bursts assembly utilizing edge node functionalities and also forward transit bursts to intermediate nodes using core node functionalities. Arriving packets from periphery user edge networks are classified according to their destination address as well as traffic class before being forwarded to assembly queues. The node uses the segmented burst assembly algorithm as well as adjustable offset timing [1], [2]. The segmented data bursts are ultimately passed on to the available burst transmission queues (BTQs) for temporary storage, while awaiting scheduling. Finally, they will be passed on to a scheduler for scheduling on available outgoing channels. Prior to scheduling, a CP is sent ahead at an offset time [2].

The same node can also handle transit data burst connections. The associated BCP of a transit data burst connection is processed in the routing module, normally availed at each node. If the received BCP is signaling a local terminating connection, then provision will be made to forward the data burst to one of the disassembly modules for its disassembling into individual data packets. However, in case of a transit connection, both the CP and associated data burst are rescheduled to the desired next node. This is subject to the necessary resources, such as the original wavelength being available. However, if contention occurs, the received data burst may have to be reticulated via the feedback unit until such time that the resources (desired wavelength) becomes available, otherwise it is discarded. Though not illustrated in the generic node architecture of Figure 1, buffer provisioning is necessary at assembly, burst transmission queues and as well as at schedulers.

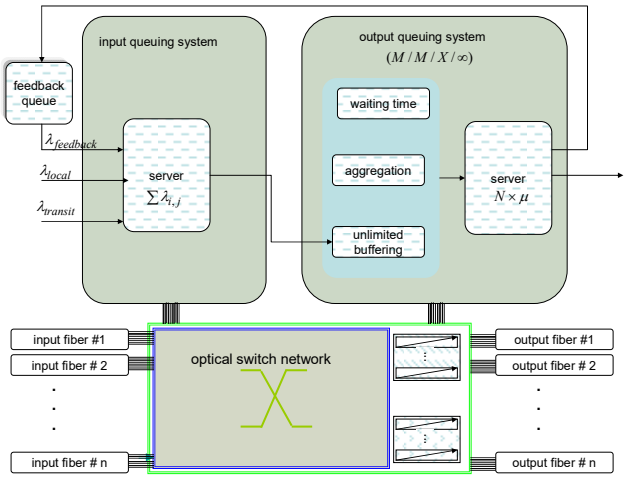

Figure 2.An Example Queuing model of an OBS Node

Whereas buffer provisioning is restricted to ingress nodes, and with none in the core nodes, however in practice most nodes are composite, i.e. they incorporate capabilities of originating, transiting as well as terminating lightpath connections. A generic queuing model of the OBS node is provided in Figure 2. Three types of connections namely, local, transit as well as feedback are served. The buffering provisioning is implemented in the form of fiber delay lines (FDLs) and flash (electronic) memory. Both can only render deterministic as well as limited delay even though it is often assumed that any burst losses are only due to wavelength contention rather than buffer overflows. Whereas traditionally we can always assume either a $M/M/X/\infty$ or $M/M/X/k$ model in terms evaluating such a model, it should however be noted that the number of input flows (streams) are limited and as such, the overall burst loss probability at an OBS link is actually lower [3]. Partly, this is because bursts within one input lightpath connection (stream)

are often streamlined and only inter-stream contentions happen at the link. This is refereed to as the streamline effect

III. RELATED WORK

The task of contention minimization in OBS switched backbone networks is accomplished by proper dimensioning of necessary and available resources at wavelength assignment, link and path levels. The key constraint being that more than one data burst cannot be assigned the same wavelength concurrently on the same link. At wavelength assignment level, various schemes such as random wavelength assignment, first-fit (FF), minimum product, maximum sum, best-fit least loaded, least utilized, most frequently used and relative capacity loss have been explored [4]. The FF scheme generally performs relatively better in terms of burst loss probability and fairness. Furthermore, it has low computational overhead and complexity. To maximize on the number of simultaneous end-to-end lightpath connections, wavelength reassignment algorithms using minimum overlap and reconfiguration techniques have been suggested [5]. However, the suggested techniques only slightly reduce the blocking probabilities. The priority-based FF offline wavelength assignment scheme proposed in 6 is geared towards maximizing both the number of simultaneous connections as well as low burst losses. With this scheme, the wavelengths to be utilized for the connection requests are prioritized according to their estimated burst loss probabilities. The priority-based FF approach requires a longer setup time as it requires extra processing time to further estimate the loss probabilities on each selected lightpath connection.

At link and path levels, it is desirable that the shortest light path(s) from ingress to egress node be utilized, subject to constraints such as traffic load, congestion as well as wavelength assignment. As suggested in [7], efficient routing can be partly achieved by ensuring that path computation is optimized as much as possible. Examples include the Dijkstra algorithm-based routing protocols such as the O*pen Shortest Path First* (OSPF) and the *Intermediate System to Intermediate System* (ISIS). Whereas they always thrive to find an optimal path for each ingress to egress node pair, they however cause the same shortest links to become congested as well as be prone to contentions. With respect to the ingress-node destination pair, the longer paths remain underutilized and overall there is traffic imbalance in the network. In order to counter this, authors in [8] propose a distributed Path Computation Element (PCE) that enables routing protocols to efficiently utilize all available network links. PCE also applies software-defined networking (SDN) paradigms to separate signaling and routing paths, thus giving more network control to operators and in that way contentions are reduced overall. An algorithm called the Self-Tuned Adaptive Routing (STAR) [9], was further incorporated to enhance traffic balancing as well prevent links from being overwhelmed.

A dynamic contention as well as congestion aware scheme that seeks to reduce blocking probabilities as well as boosting utilization by symmetrically distributing network traffic over all active links was proposed in [10]. Finally, in [11], the researchers proposed and investigated a per-link congestion control-based scheme that seeks to balance available network resources allocation by utilizing present and forecast demands of lightpath requests statistics. In essence, practical networks have a regularized topology and lightpath connection requests are generally random in nature. Given a fixed amount of resources (link, wavelengths, paths, as well as constraints), an increase in the traffic load results in the reduction of the number of idle resources per link and hence this will lead to both contention as well as blockings.

We propose a priority-based limited intermediate buffering and streamline effect aware prioritized routing and wavelength assignment (RWA) scheme (LIB-PRWA) to combat the problem of contention occurrences. The approach relies on prioritized grooming of local and transit lightpath connection requests. This is followed by prioritizing wavelengths according to their past performances in terms of contention occurrences on each. Finally, it assigns the wavelength to the various connection requests by further taking into consideration other resources states (such as congestion, and current traffic loads) in primary paths. Summarily, the paper's contributions are as follows:

a) We introduce a burst grooming algorithm for mixing of transit and local data bursts at core nodes. As discussed later, the grooming helps in minimizing contentions.

b) We propose and discuss a limited intermediate buffering and RWA (LIB-PRWA) based scheme in which contending bursts may be buffered at a core node depending on their residual hop count. As such, it will be shown that this helps to improve the fairness in terms of drop rate of different hop-count bursts.

The rest of the paper is organized as follows: A short discussion on streamline effect aware RWA as well as constraints is provided in section IV. The proposed scheme as well as its key elements such as Priority based RWA, Traffic Grooming and a limited buffering node architecture model are narrated in section V. The scheme's performance analysis is provided in section VI, thereafter section VII concludes the paper.

IV. STREAMLINE EFFECT AWARE RWA

The extent to which the streamline affects the overall burst loss ratio in the network for individual as well as aggregated flows was explored in various literatures e.g [12]. Overall, at an arbitrary node, the burst loss ratio of all data bursts constituting a stream on an individual end-to-end lightpath connection between a source (u_s) and destination (u_d) pair is the same, and contentions are rather among the various streams on the link. It has also been shown that both the burst loss ratio as well as contention become lower as the traffic arrival rate of a flow increases.

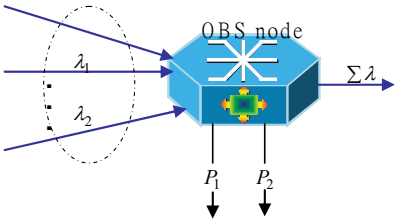

Figure 3. Flow aggregation at a Node

On the basis of the aforesaid, for a given OBS network, $G(V,L)$ comprising V nodes and a total of L links (fibers), the objective would be to either maximize the simultaneous supported network traffic $D_{s,d}$, or minimize unsupported traffic $U_{s,d}$ for each node pair $(v_s,v_d)\in S,D$. The main objective of a SLE aware RWA scheme would be to maximize $D_{s,d}$ subject to optimized network supporting resources.

$$\max \sum_{(v_s,v_d)\in SD} D_{s,d} \tag{1}$$

Alternatively the objective function is:

$$\min \sum_{(v_s,v_d)\in SD} U_{s,d} \tag{2}$$

So, SLE aware RWA will strive to groom as much traffic as can be possible within a single wavelength of capacity B.

$$\sum_{l\in\omega^+(v)} \varphi_{l,\lambda}^{s,d} = \begin{cases} b_\lambda^{s,d} & \text{if } v = v_s \\ -b_\lambda^{s,d} & \text{if } v = v_d \\ 0, & \text{otherwise} \end{cases} \quad v\in V,(v_s,v_d)\in S,D \tag{3}$$

$$D_{s,d} = \sum_{\lambda\in\Lambda} b_\lambda^{s,d} + U_{s,d} \qquad (v_s,v\in \mathrm{S,D}) \tag{4}$$

$$\sum_{(v_s,v_d)S,D} \varphi_{l,\lambda}^{s,d} \le B \quad l\in L,\ \lambda\in\lambda \tag{5}$$

The above two equations assume that each lightpath request $D_{s,d}$ is served on a single flow path as well as wavelength. Also note that for $\lambda\in\lambda,(v_s,v_d)\in S,D$;

$$b_{\min} y_{l,\lambda}^{s,d} \le \varphi_{l,\lambda}^{s,d} \le D_{s,d} y_{l,\lambda}^{s,d} \quad l\in L \tag{6}$$

$$\sum_{l\in\omega^+(v)} y_{l,\lambda}^{s,d} \le 1;\ \sum_{l\in\omega^-(v)} y_{l,\lambda}^{s,d} \le 1; \quad v\notin\{v_s,v_d\} \tag{7}$$

$$\sum_{l\in\omega^-(v_s)} y_{l,\lambda}^{s,d} = \sum_{l\in\omega^+(v_d)} y_{l,\lambda}^{s,d} = 0 \tag{8}$$

$$\sum_{l\in\omega^+(v_s)} y_{l,\lambda}^{s,d} = \sum_{l\in\omega^-(v_d)} y_{l,\lambda}^{s,d} \le 1 \tag{9}$$

The following constraints will ensure the SLE;

$$y_{l_i,l_o,\lambda} \ge y_{l_i,\lambda}^{s,d} - 1 \ \lambda\in\lambda,(\mathrm{v_s},v_d)\in S,D; v\in V\setminus\{v_s,v_d\} \quad l_i,l_o\in L:\ \mathrm{DST}(l_i)=\mathrm{SRC}(l_o)=\{v\} \tag{10}$$

$$\sum_{l_i\in L:\,\mathrm{DST}(l_i)=\mathrm{SRC}(l_o)} y_{l_i,l_o\lambda} \le 1,\ l_o\in L,\ \lambda\in\Lambda \tag{11}$$

The last two equations are also an indicator of wavelength connectivity (continuity) at each node, hence we have the following variables:

$$y_{l,\lambda}^{s,d} \in \{0,1\} \qquad l\in L,\ (v_s,v_d)\in S,D;\ \lambda\in\Lambda \tag{12}$$

$$y_{l,\lambda}^{s,d} \in \{0,1\} \qquad l_i,l_o\in L:\ \mathrm{DST}(l_i)=\mathrm{SRC}(l_o),\ \lambda\in\Lambda \tag{13}$$

$$\varphi_{l,\lambda}^{s,d} \ge 0 \quad l\in L,\ \lambda\in\Lambda,\ (v_s,v_d)\in S,D \tag{14}$$

$$b_{l,\lambda}^{s,d} \ge 0 \quad \lambda\in\Lambda,\ (v_s,v_d)\in S,D \tag{15}$$

$$U_{s,d} \ge 0 \quad \{s,d\}\in S,D \tag{16}$$

In the last five sets of equations, all the variables are indexed with respect to the wavelength λ. In order to reduce the number of variables required to solve each set of SLE aware RWA constraints, a decomposition model approach can be used. The model partly utilizes parameters and variables obtained using a the column generation model approach, which itself relies on a set of dynamically generated routing and provisioning network configurations. Each SLE compliant carries a fraction of the $D_{s,d}$ traffic for a designated source and destination node pair $(v_s,v_d)\in S,D$. Specifically, the SLE-RWA decomposition model relies on key parameters such as:

S,D – set of source node pairs (v_s,v_d), with $D_{s,d}>0$

C – set of wavelength configurations

c – generic wavelength transmission , which is characterized by a set of SLE compliant paths (p^c).

$$p_l^c = \begin{cases} 1, & \text{if link } l \text{ belongs to path} \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

$b_{s,d}^c$ – bandwidth used in c for $(v_s, v_d) \in S, D$.

$a_{s,d}^c = 1$, if there exists at least one $v_s \to v_d$ flow in configuration c .

z_c – is a variable denoting the number of wavelength for which configuration c is selected.

We can rewrite equation (1) as follows:

$$\min \sum_{(v_s, v_d)} U_{s,d} + \text{PENAL}_W \times \sum_{c \in C} z_c + \text{PENAL}_p \times \times \sum_{c \in C} \left(\sum_{(v_s, v_d) \in S, D} a_{s,d}^c \right) z_c \tag{18}$$

Equation (18) has three parts , the first minimizes the amount of available bandwidth not utilized, the second minimizes the number of utilized wavelengths, and the third maximizes the bandwidth usage.

This is subject to the following constraints;

$$\sum_{c \in C} z_c \le W \tag{19}$$

$$\sum_{c \in C} d_{s,d}^c z_c \ge D_{s,d} - U_{s,d} \quad (v_s, v_d) \in S, D \tag{20}$$

$$z_c \in Z^+ \quad c \in C \tag{21}$$

Equation (19) imposes a usage limit on the number of available wavelengths, whereas equation (20) is a demand constraint subject to (19).

V. PROPOSED SCHEME

Before a data burst is dispatched, resources have to be provisioned for it. This involves determining the least cost path to destination node on which the lighpath connection will be established, and assigning a wavelength to it. It is possible to avoid contention occurences at the next and subsequent nodes by either assigning different links or different wavelength to concurrent bursts originating from the same node. However, at the next and subsequent nodes, each light path connection (data burst) is likely to merge with other transit connections. In so doing on any link, different wavelengths must be assigned to each of the bursts to avoid possible contentions, should they partially or wholly overlap in time. The SLE-aware based RWA's goal is to maximise the number of simultaneous lightpath connections establishments to various source – destination node pairs within the network subject to these constraints. The contention resolution mechanisms at nodes must not escalate network costs, degrade peformance (due to losses), or worsen contentions and other network peformance metrics in certain sections of the network. In certain instances, a data burst finds itself being discarded when it is only a few hops from the destination node, and this would be quite wasteful of resources. The proposed scheme involves enforcing a few measures such as traffic grooming at nodes, selection of both shortest possible paths to destination followed by the evaluation of their current resources states. The selected routes will be prioritized according to the frequency of contention occurrences as well as current network resources metrics. Limited buffering to contending bursts in the form of a feedback unit incoporated in each core node is also implemented.

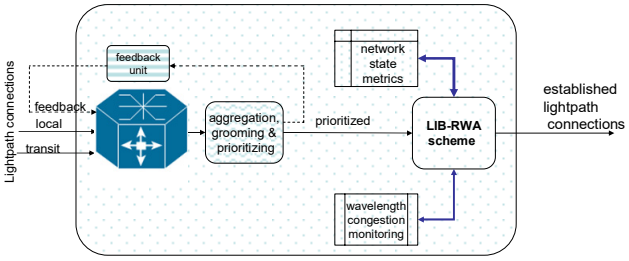

Figure 4. Proposed scheme's concept

As illustrated in Figure 4, lightpath connection requests are aggregated and then groomed according to priority (low or high priority). After grooming, the LIB-PRWA scheme will choose routes (including deflection routes) based on current network state as well as frequency of contention occurrences. In the process, contentions can also be resolved by reticulating one or more of the contending data bursts via the incorporated feedback unit. The various steps are summarily discussed next:

A. Aggregation

The SLE aware RWA is one of the key components of the proposed scheme responsible for aggregating traffic both at burst as well as lightpath levels at the input ports. The traffic includes that generated by local (originating), transit as well as feedback (rearticulated) bursts. The aggregation utilizes the available aggregation buffers in serving mostly incoming local and feedback traffic. The traffic (bursts) are placed in each aggregation queue according to destination, priority as well as distance. This means bursts destined for a common destination node are served in the same queue. Feedback as well as other incoming high priority bursts will always have precedence over the rest. As part of the SLE awareness, bursts destined for shorter hops are accommodated ahead of those traversing longer distances. Time wise, this implies that when the aggregated data burst arrives at the next node, sections destined for shorter distances (including this node) are disaggregated and the now vacant window made available for new traffic aggregation as well as accommodation of any feedback traffic at this node. Overall, the process generally is designed to ensure the seclusion of any

overlapping with all other incoming transit traffic bursts. In practice, secondary contention may still occur. An example is when the data burst aggregation is already planned and near completion, and the new overlapping aggregated data burst's associated BCP arrives.

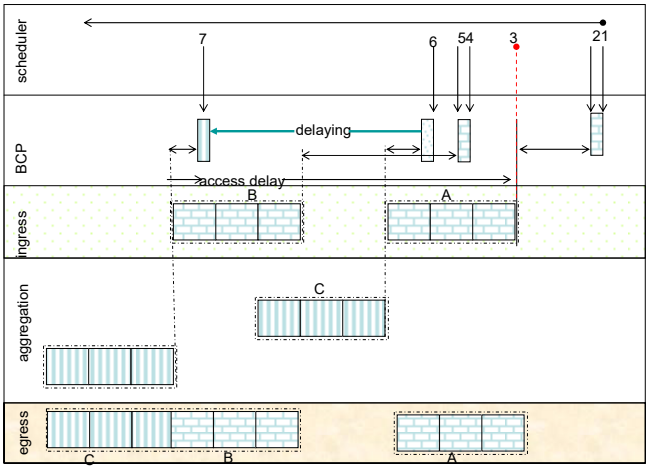

Figure 5. An example of resolving secondary contention

This normally will happen in cases where the incoming aggregated burst is very close to its destination, hence the offset time is not enough such that the aggregated burst arrives almost at the same time as the BCP. As a remedy, an adjustable scheduling mechanism that has an allowance for aggregation/ unbuffering delaying is implemented. This is illustrated in Figure 5.

Algorithm I: SLE Aware Rescheduling

initialize

1. Incoming BCP is registered for processing
2. $p \leftarrow$ i/c port of i/c aggregated data burst
3. $\lambda \leftarrow$ wavelength of i/ aggregated burst
4. $t_{arr} \leftarrow$ i/c aggregated data burst arrival time
5. $\Delta \leftarrow$ aggregated data burst's length
6. check_if_contention (p, λ, t_{arr}) **then**
7. *CanReschedule* $\leftarrow$ **if _true**
8. **for** *Overlapping Scheduling* $s_p^{\lambda} \rightarrow$ s , **do**
9. **if** BCP of s already dispatched **then**
10. *CanReschedule* $\leftarrow$ **if _False**
11. **if** *rescheduling possible* **then**
12. *reservation* of λ from t_{arr} for Δ
13. **for** *scheduling* $s_p^{\lambda} \rightarrow$ s , **do**
14. RecomputeWaiting (s)
15. **else**
16. aggregated burst is reticulated
17. else
18. *reserve* of λ from t_{arr} for Δ

end

B. Grooming

Primarily, the purpose of grooming the connection requests is to improve on network utilization. The groomed connection requests are further prioritized so that precedence is given to requests with relatively higher priorities. In that way, more requests are likely to be successfully established simultaneously in the process. Contentions as well as blocking probabilities are drastically reduced as a result of the grooming and wavelength prioritization facilitated by the proposed LIB- PRWA scheme. A summary of a priority grooming algorithm is as follows [13], [14].

Algorithm II: grooming and prioritizing

initialize

Step1: lightpath connection requests destined for the same (s,d) groomed within link capacity

$$R = \left\{ r_1^{s,d}, r_2^{s,d}, ..., r_n^{s,d} \right\} \Big/ \sum_i^n B(r_i^{s,d}) = gr^{s,d}$$

Step II: queue all lightpath connection requests.
Step III: categorize them into low and high priority.
end

C. Overall link States

a) **Link/Path Congestion**: We commence with link as well as path congestion. To this end, we utilize the maximum link threshold value $\rho_{\max}^{s,d}$ and if the link threshold value exceeds the set threshold value, the link state (LS) is set to 1 .

$$\rho_{i,j} \geq \rho_{\max}^{i,j} = LS_{i,j} = \begin{cases} 1 \rightarrow congestion \\ 0 \rightarrow available \end{cases} \tag{21}$$

The path congestion state on the primary path is determined from:

$$LS_{s,d}^{p} \underset{LS \neq 1}{=} \sum_{e(i,j) \in r^{p(s,d)}} LS_{(i,j)} \tag{22}$$

The proposed scheme will always give preference to paths with high connection establishment success likelihood, i.e paths in which congestion likelihood is at a minimal ($\min(LS_{(s,d)}^{p})$). The same applies to links. In weighted terms, congestion levels at any given time t can be computed from;

$$c_{(i,j)}(t) = \frac{N_{drop(i,j)}}{N_{total(i,j)}} \tag{23}$$

Assuming link blocking probabilities to be independent, then at any time, the overall blocking probability is:

$$P_{B(\pi_k,t)} = 1 - \prod_{1 \leq i \leq n_k} (1 - c_{(i,j0)}) \tag{24}$$

b) **Wavelength utilization**: With regards to wavelength utilization, it is generally noted that data bursts routed on paths and links that are least used are not likely to encounter any contention. Furthermore, in the unlikely event that contention occurs along the path, the limited available contention resolution mechanisms will suffice to prevent any burst discarding.

At any given time, the utilization of a link is determined from [14].

$$U(e,t) = \frac{\sum_{i=1}^{N_s} T_i^S}{W_l \times t} \tag{25}$$

where W_l denotes the total number of usable (active) wavelengths along a given link.

c) **Wavelength congestion monitoring:** This is a tracking of the frequency of usage of all individual active wavelengths. The statistics of contentions recorded for each of the wavelengths is also recorded. This information is normally acquired at fixed time intervals Δ for each output link $(f_c^{i,j})$ throughout the network and periodically shared with all other nodes.

d) **Intermediate Buffering:** If a channel scheduler is unable to find a free desired wavelength on the next outgoing link, the data burst is discarded. In the proposed scheme, we assume that each core node incorporates an intermediate buffering provisioning to cater for those data bursts that are nearing their respective destinations. The buffering is implemented in the form of a feedback unit that incorporates FDLs at each node. Typically, data bursts that have traversed half the network's radius $(\Re_{s,d})$ are eligible for intermediate buffering. We define the network's radius as half the maximum hop count between the longest of the shortest paths possible in a given network. As argued in various literatures, e.g. in [15], lightpath connections serving a source-node destination pair spanning long hop counts are quite likely to encounter contentions and that their discarding (as a contention resolution measure) may adversely affect the overall network throughput. Note that the discarded bursts would have already utilized significant amounts of network resources.

Algorithm III: LIB-PRWA

initialize
input: acquire sets of new and transit connection requests from CP processing module.
output: sets of lightpath connection requests; These are classified as low or high priority.
Step I: acquire network metrics: congestion level, contention frequency, utilization and search for K shortest paths search for set of shortest paths,
Step II: Serve all requests according to priority.
Step III: From fail list: transit connection request check hop distance $(if \geq \Re_{s,d})$ **, send to IBQ.** Re-set priority to highest, and repeat step II once
Step IV: drop any fails
end

D. Limited Buffering at Nodes

The generic architectural core node block diagram of Figure 1 incorporates a feedback unit that facilitates limited buffering. Its functional queuing model equivalent as illustrated in Figure 6 would be two nodes A and B representing the core node and its incorporated feedback unit respectively.

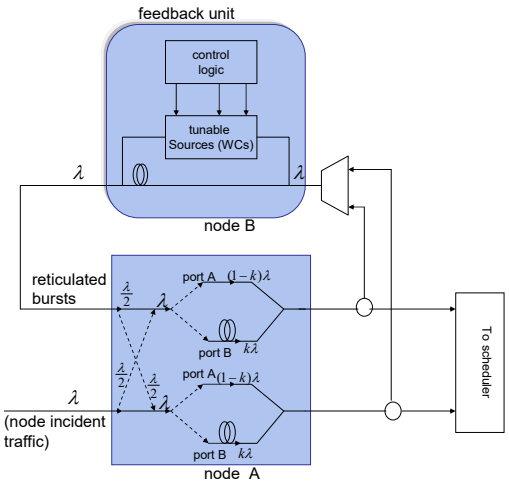

Figure 6. Model of the proposed feedback unit providing limited buffering

Node A comprises an $n \times n$ optical switching fabric, where n - is the number of optical links at the input and output respectively. Node B represents the incorporated feedback unit that provides limited buffering and generates a traffic load λ representing the contending burst that were looped back. As illustrated, node A prior to grooming, provides two ports A and B. In case of contention, one or more of the contending bursts from both ports A and B are looped back via the feedback unit. The probability of sending an arriving burst to port A is k, whereas that of sending it to port B is $1-k$ queues. In order to determine the overall node blocking probability P_N (i.e. taking into account the feedback unit) we proceed as follows;
We define the probability that port A is busy as ;

$$B_1 = \frac{k\lambda}{k\lambda + \mu} + \frac{(1-k)\lambda}{(1-k)\lambda + \mu} \tag{26}$$

The probability that port B is busy is ;

$$B_2 = \frac{k\lambda}{k\lambda + \mu} \tag{27}$$

The probability (B_3) of bursts existing in the delay of port B is;

$$B_3 = \left[1 - \frac{\mu}{\mu + \left(\frac{k\lambda\mu}{k\lambda + \mu} \right)} \exp - \left(\frac{k\lambda\mu}{k\lambda + \mu} T \right) \right] \tag{28}$$

where T denotes the momentary delay in the feedback unit. A burst will find port B busy with a probability;

$$B_4 = \left[1 - \frac{\mu}{\mu + \left(\frac{k\lambda\mu}{k\lambda+\mu}\right)} \exp - \left(\frac{k\lambda\mu}{k\lambda+\mu} T\right)\right] \times \left[\frac{k\lambda}{k\lambda+\mu} + \frac{(1-k)\lambda)}{(1-k)\lambda+\mu}\right] \quad (29)$$

Since the overall performance of any network is quantified taking into account factors such as the utilization (U), bandwidth (B`) as well as link rate (R), we can thus define a throughput characterization factor;

$$\Gamma = [0,1] = \exp\left(\frac{-R}{UB}\right) \quad (30)$$

From which the blocking probability of the node (excluding the feedback unit) is

$$P' = \exp\left(-\frac{R}{U \cdot B}\right)\left[\left(1 - \frac{\mu}{\mu + \frac{\lambda\mu}{+\lambda+\mu}}\right)\exp - \left(\frac{k\lambda\mu}{k\lambda+\mu} T\right)\right] + \left[\frac{k\lambda}{+\lambda+\mu} + \frac{(1-k\lambda)}{(1-k)\lambda+\mu}\right]\left[\frac{\mu}{\mu + \frac{k\lambda}{k\lambda+\mu}}\exp - \left(\left(\frac{\lambda\mu}{\lambda+\mu}\right)\right)\right] \quad (31)$$

The overall joint probability blocking of the node with feedback unit can be determined by first assuming that all bursts are of equal length, the burst arrival process at all inputs follows a Poisson distribution and that the node can handle a maximum of B_N bursts at any given time. In this case we first determine the average numbers of bursts at the node as well as average waiting time in the feedback unit. Consequently, we get;

$$P_N = \exp\left(-\frac{R}{U \cdot B}\right)\left[\left(1 - \frac{\mu}{\mu + \frac{\lambda\mu}{+\lambda+\mu}}\right)\exp - \left(\frac{k\lambda\mu}{k\lambda+\mu} D\right)\right] + \left[\frac{k\lambda}{+\lambda+\mu} + \frac{(1-k\lambda)}{(1-k)\lambda+\mu}\right]\left[\frac{\mu}{\mu + \frac{k\lambda}{k\lambda+\mu}}\exp - \left(\left(\frac{\lambda\mu}{\lambda+\mu}\right)\right)\right] \times \left[\left(\frac{\lambda}{\eta}\right)^3 \frac{\left(1 - \left(\frac{\lambda}{\mu}\right)\right)}{\left(1 - \left(\frac{\lambda}{\mu}\right)^{B_N+1}\right)}\right] \quad (32)$$

According to [16], as long as the service discipline at the node is work conservative, the fraction of bursts that are blocked is independent of the service discipline.

VI. PERFORMANCE EVALUATION

We commence this section by briefly evaluating, SLE-aware RWA's performance as discussed in section IV. Performance measures of interest will include burst buffering probabilities, secondary contention as well as buffering and access delays. This will be followed by a direct performance evaluation of the overall LIB-RWA scheme in two separate cases; when intermediate buffering is provisioned and when it is not.

A. SLE Aware Aggregation Performance

As already discussed in previous sections, SLE Aware RWA aggregates bursts from various flows in such a manner as to reduce secondary (inter stream) contention. The buffering implemented in the form of feedback units does burst loss avoidance, even though this is at the expense of extra delays experienced by each buffered flow. The evaluation is carried out on a multi-node network comprising 11 nodes interconnected by 26 bidirectional links. The distance between nodes is in km. Each node incorporates a feedback unit that can provide varying delays for bursts as desired. The network is implemented in OMNET++ (version 5.4). The platform also includes OBS modules that implement both ingress and egress nodes. The ingress nodes generate and supply constant size data packets whose payloads are fixed at 100 kB. Their inter-arrival times follow a Poisson distribution. The data packets create streams that feed to aggregation queues for the generation of data bursts. The Just Enough Time (JET) principle is chosen as the signaling protocol among the various nodes, whereas burst assembling utilizes the in-built LAUC-VF algorithm. Each wavelength operates at either 10 or 100 GBps .

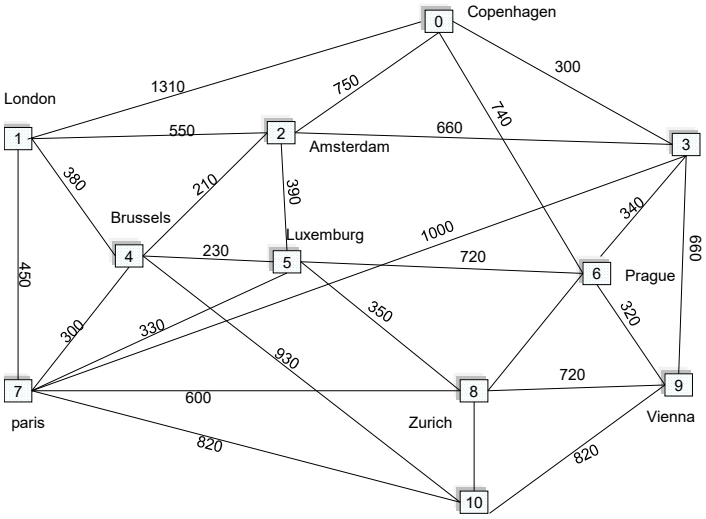

Figure 7. The Network with 11 node and 26 bidirectional links

Table 1. Simulation parameters

parameter	value
Number of nodes	11, edge-core
Edge nodes	0,1,3,7,10
path distance	Variable
path bandwidth	10Gbps/ 100 Gbps
Ave. burst length	3MB
BCP processing time	10 μs
Link status update	Every, 2 ms
Offset time	Variable =f(h)

The traffic load intensity is varied from zero to 1 and calculated as;

$$\rho = \left\lfloor \frac{S_{bursts}(t)}{B_\lambda \times W \times L \times time} \right\rfloor \quad (33)$$

where $S_{bursts}(t)$ is the aggregate size of bursts sent throughout the network, B_λ single wavelength's capacity, W is the number of wavelengths in a single link, and L is the number of links in the network.

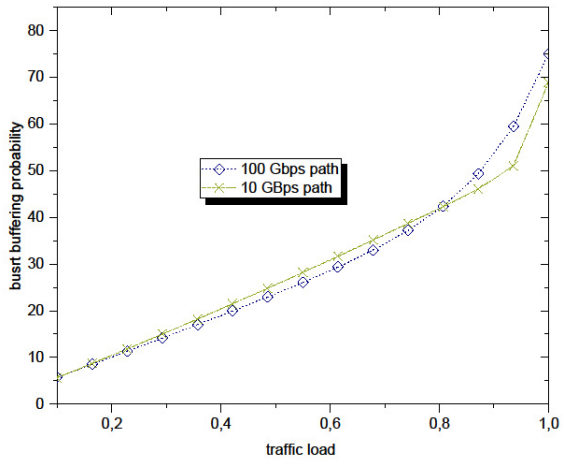

Figure 8. Burst buffering probability

Figure 8 plots the burst buffering probability as a function of traffic load. In general, it is noted that a burst flow has a high likelihood of being buffered when it is traversing a relatively long hop distance. As the overall traffic increases so is the more instances of secondary contentions which can only be resolved by temporarily buffering the affected flow (see Figure 5, for illustration purposes) till such time that desired aggregation can take place. As traffic load increases, so will be the number of secondary contentions also increase and hence more buffering likelihoods. When operating the links and paths at 100GBps, the number of flows are likely to correspondingly increase and hence the number of potential secondary contentions that would require buffering of the individual flows.

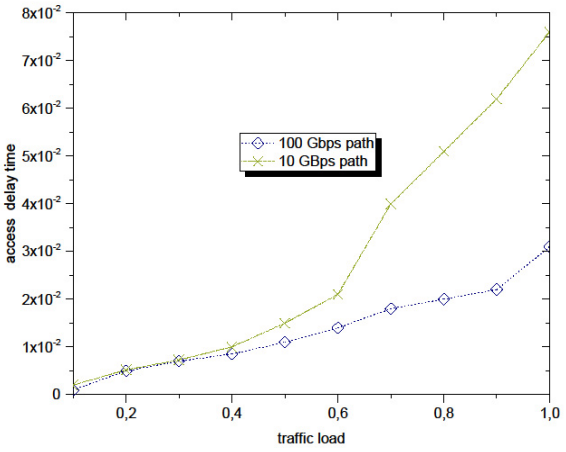

Figure 9. Access delay times

Variation of access delay times for awaiting data bursts along transit nodes are plotted in Figure 9. It is noted that at low network traffic loads, as expected, access delay times are quite small since there is lots of voids in transiting flows affording aggregation to take place. However, as the traffic load peaks above (50%) access delay times significantly increase for paths/links operating at low speeds. This is because in this case, all transiting flows tend to be filled up and hence not much voids are available to facilitate aggregation of awaiting bursts at intermediate nodes.

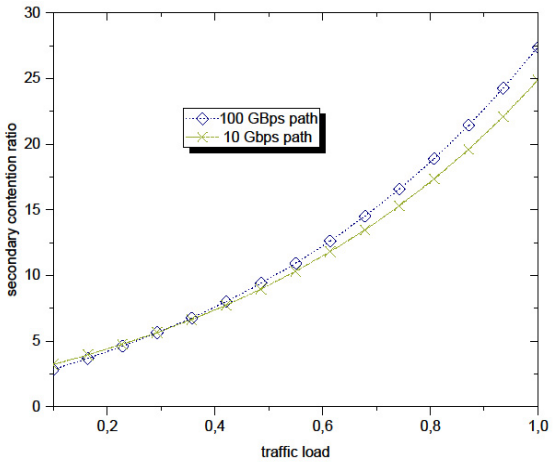

Figure 10. Secondary contention ratio

Secondary contention ratios are plotted as a function of traffic load when operating the networks at 10 GBps and 100 GBps respectively. Generally, the magnitude of such contention flows is an indicator of the extent of aggregation at nodes since SLE aware RWA tends to avoid primary contention. As can be

observed for both speeds, when traffic is increased the algorithm uses more wavelengths hence the latter's more efficient use and thus leading to higher link utilization throughout the network.

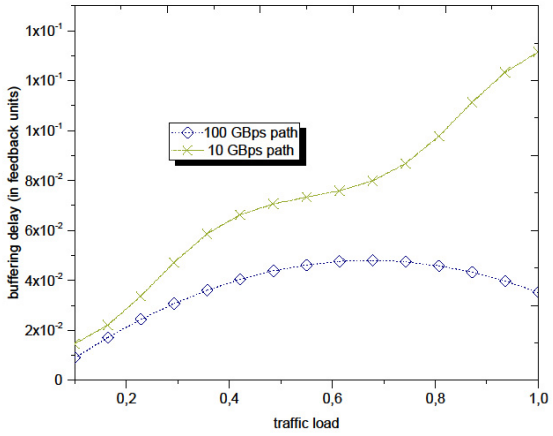

Figure 11. Nodal buffering delays

One of the distinct features of SLE aware RWA is that of minimizing nodal buffering delays at all ingress nodes. As can be observed from Figure 12, when operating at 100 GBps, minimum buffering delays are incurred at both low and high network traffic loads . This is because at low traffic loads, the voids will always be available for the aggregating of awaiting bursts at the nodal nodes on transit flows. Operating the network at higher speeds will mean more wavelengths as well as voids are available and hence the nodal buffering delays are quite low as well.

B. LIB-PRWA without Intermediate Buffering

In this subsection, we carry out a performance comparison of our proposed scheme versus the already existing routing ones. We assume no intermediate buffering, i.e. the incorporated feedback units on each core node are assumed offline.

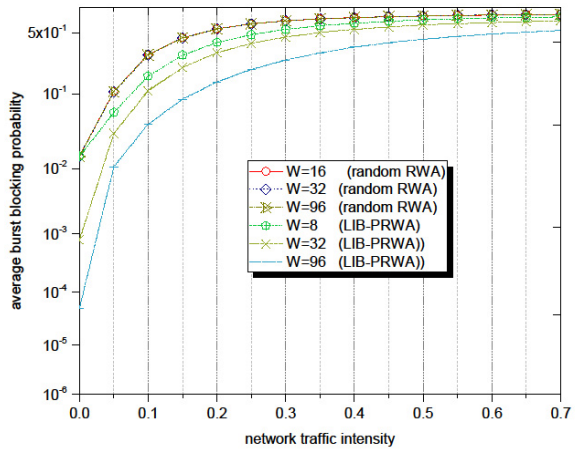

Figure 12. Blocking probabilities as a function of traffic load

We compare the performance of the proposed scheme versus random RWA (representing the traditional OBS approach). The number of wavelengths per link is gradually increased. The blocking performance is plotted for values of $W = 16, 32$ and 96 in Figure 13. As can be observed, the blocking performance in the case of random RWA is more or else identical when the number of wavelengths is varied . However, the proposed scheme's performance is relatively significant as the number of wavelengths per link is increased.

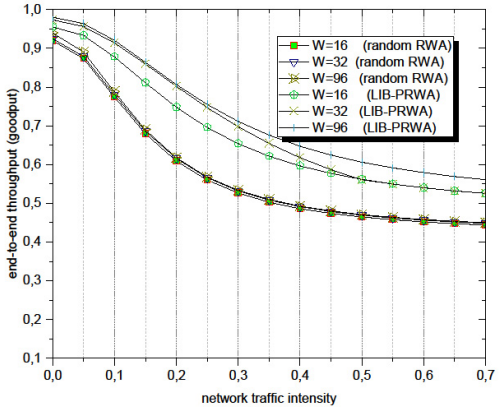

Figure 13. Network throughput

Figure 14 plots the overall end-to-end network throughput. The random RWA scheme displays the same trend with regards to the throughput for various values of W .

Overall, the proposed LIB-PRWA approach enhances network performance by reducing blocking and at the same time increasing the throughput. As expected, random RWA algorithms' performances for varying wavelength capacities are identical. On the other hand, the LIB-PRWA algorithm outperforms the random RWA quite distinctively. It is also noted that for low values of W , the limited resources rather dictate blocking probability and not the wavelength assignment approaches implemented.

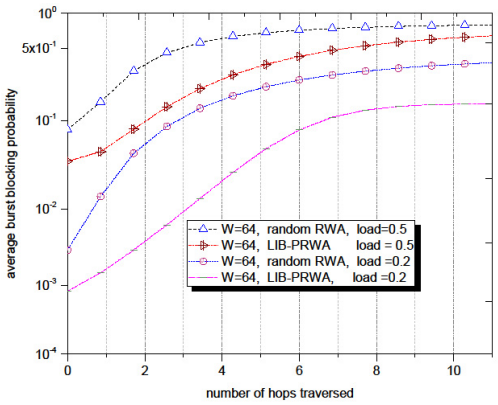

Figure 14. Blocking as a function of hop count

The overall performance improvement of the LIB-PRWA with increases in W can be attributed to the degree of wavelength spatial reuse, i.e. for large values of W, an ingress node can schedule more lightpath connections (bursts) on a given link. Consequently, more lightpath connections traversing different links can be concurrently assigned the same wavelength values. Furthermore, by comparing the two schemes at low traffic levels, the LIB-PRWA has relatively better performance. This is because wavelength contentions prominently contribute to the blockings at low traffic loads, whereas as the network traffic load increases, most of the burst blockings are also caused by insufficient bandwidth. With regards to the number of nodes (hops) traversed, we note that for low traffic loads, the LIB-PRWA algorithm improves the network performance in terms of the blocking. As can be observed from Figure 15, the two approaches more or less perform identically at high loads, indicating that no more wavelengths are available for newly generated bursts and senders have to block them immediately.

C. LIB-PRWA with Intermediate Buffering

In this subsection, we compare the performance of the proposed scheme when it enforces intermediate buffering. This implies that the feedback unit is functional We set the network diameter to 8. Figure 16 plots the average burst blocking probability as the link load is gradually increased. As anticipated, the LIB-PRWA performs better than the other two. Overall, it is noted that intermediate buffering is effective in contention resolution and consequently improving network performance in terms of blocking probabilities as well as improving fairness to those data bursts that traverse the network through high hop counts.

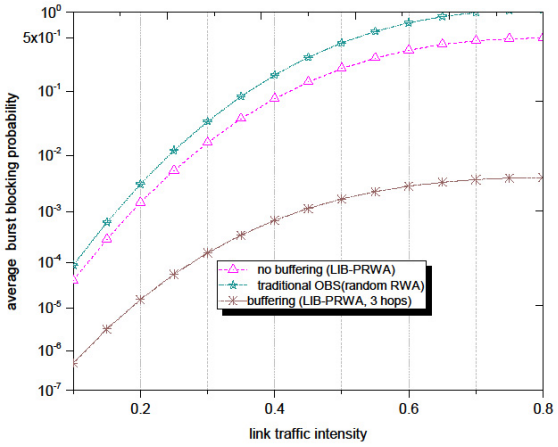

Figure 15. Average blocking capacity when link traffic is increased

The burst loss performance is compared for the three schemes namely, the proposed scheme, the traditional OBS routing's SPF (which uses random RWA) and SPDR in Figure 17.

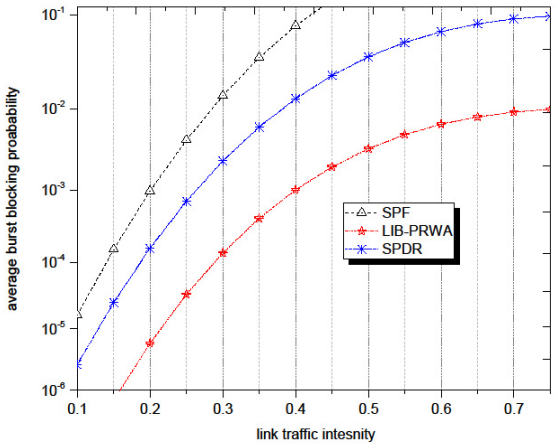

Figure 16. Average blocking probability SPF, LIB-PRWA and SPDR

The average blocking performance for the three schemes is plotted as a function of the offered link traffic intensity (load). Beyond load values of 0.4, the proposed scheme's burst loss continues to be relatively low. This is because as the traffic increases, the feedback unit can no longer accommodate all the reticulated data bursts. Further, if we define the coefficient of variation (an indicator of the degree of unfairness to individual traffic connections in the network) to be the ration of the standard deviation (σ) to the mean (μ), then the traditional scheme performs relatively better when network loading conditions are below 0.7, whereas the proposed scheme performs relatively better in high network loaded scenarios. Figure 18 provides a plot of the coefficient of variation of the blocking probability as the link load is varied steadily from 0.2 to about 0.8.

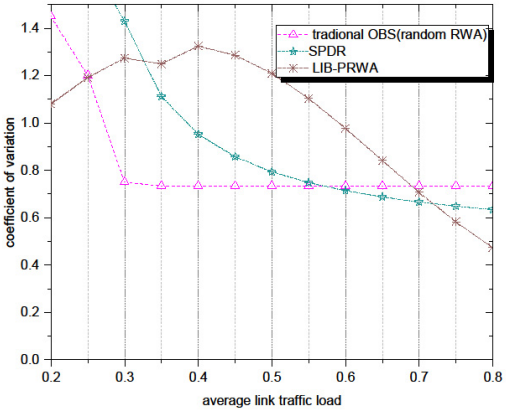

Figure 17. Coefficient of variation of the blocking probability

The proposed scheme performs relatively better than the rest. Figure 19 plots the end-to-end throughput for selected routing strategies considering relatively uniform as well as distance-dependent traffic.

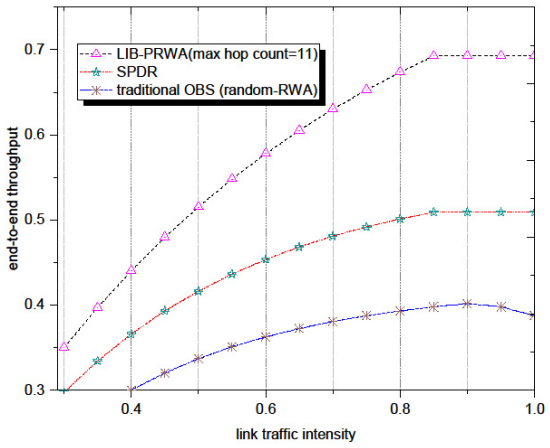

Figure 18. End to end throughput as a function of link load

Both SPDR and the proposed scheme outperform SPR. However, the proposed scheme utilizes the available network resources much more efficiently and shows the highest throughput overall.

VII. CONCLUSIONS

The paper addresses the problem of contention occurrences in OBS networks. Frequent contention occurrences in such networks will lead to high data burst losses and consequently a degrading of QoS for the running applications and services. Primarily, in this paper we distinguish two types of contention, primary and secondary. Primary contention is caused by two or more data burses contending for the same output port wavelength at the same time, whereas secondary contention will register when the unbuffering of a previously temporarily buffered data burst flow in the feedback unit overlaps in time with that of another incoming burst when they are requesting the same wavelength. Limited buffering coupled with SLE aware prioritized RWA based scheme is proposed. The limited buffering at all intermediate nodes caters for those data bursts that have traversed more than half the network's radius and suddenly encounter contention. In that way, discarding them as a consequence contention resolving would lead to low network throughput. The SLE aware RWA has been shown to seclude primary contention and aids in the proper aggregation processes at intermediate nodes thus ultimately avoiding secondary contention occurrences A queuing analysis of a typical node taking into account local, transit as well as feedback traffic is carried out. Performance of the model in terms of burst buffering probabilities, access as well as nodal delays is carried out. It is generally found out that bursts that traverse longer hops have a higher probabilities of being buffered at intermediate nodes as they are greater chances of encountering contentions. Nodal delays for data bursts awaiting aggregation at intermediates nodes can be minimized by operating the network at higher speeds as this provisions more wavelengths as well as voids for the aggregation. The proposed LIB-PRWA overall has better performance in comparison to other equivalent ones. This includes improved end-to-end blocking probabilities, throughput as well as overall network resources utilization.

.

REFERENCES

[1] H. Sibanda and B. Nleya, "Model Analysis of Dynamic Priority Segmented Burstification and Deflection Routing", Proceedings of IEEE's Africon'17, Cape Town, 18-20 September, 2017.

[2] R. Vallejos, A. Beghell,V.Albornov and M. Tafifeno. "Joint routing and dimensioning of optical burst switching networks", Photonic Network Communications volume 17, pp s266–276, 2009

[3] L. Battestilli, H. Perros and S. Chukova, "Burst lost probabilities in a queuing network with simultaneous resource possession: a single-node decomposition approach," in *IET Communications*, vol. 3, no. 3, pp. 441-453, March 2009.

[4] Routing and Wavelength Assignment. In: Optical WDM Networks. Optical Networks. Springer, Boston, MA, 2006, https://doi.org/10.1007/0-387-29188-1_7.

[5] P Chu, T. Bu and X. -. Li, "A Study of Lightpath Rerouting Schemes in Wavelength-Routed WDM Networks," *2007 IEEE International Conference on Communications*, Glasgow, 2007, pp. 2400-2405.

[6] B. C. Chatterjee, N. Sarma and P. P. Sahu, "Priority based routing and wavelength assignment with traffic grooming for optical networks," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 4, no. 6, pp. 480-489, June 2012.

[7] K. G. Ramakrishnan and M. A. Rodrigues, "Optimal routing in shortest-path data networks," in *Bell Labs Technical Journal*, vol. 6, no. 1, pp. 117-138, Jan.-June 2001.

[8] F. Paolucci, F. Cugini, A. Giorgetti, N. Sambo and P. Castoldi, "A Survey on the Path Computation Element (PCE) Architecture," in *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1819-1841, Fourth Quarter 2013.

[9] N. Jain, M. Dahlin, Y.Yin Zhang, D.Kit, P. Mahajan and P.Yalagandula "Self-Tuning Aggregation for Scalable Monitoring", In Proceedings of Jain2007STARSA, 2007.

[10] R. Sankar, B. Ashok and B. Sahoo and K. Turuk, "QoS aware routing and wavelength allocation in optical burst switching networks using differential evolution optimization", Digital Communications and Networks, Volume 4, Issue 1, February 2018, Pages 3-12.

[11] L. Bhatia, G. Randeep and K Lakshman, " Fast network re-optimization schemes for MPLS and optical networks. Computer Networks", Volume 50., pp 317-331, 2006, 10.1016/j.comnet.2005.03.004.

[12] M. H. Phung, D. Shan, K. C. Chua and Gurusamy Mohan, "Performance analysis of a bufferless OBS node considering the streamline effect," in *IEEE Communications Letters*, vol. 10, no. 4, pp. 293-295, April 2006.

[13] A. Celik, A. AlGhadhban, B. Shihada and M. Alouini, "Design and Provision of Traffic Grooming for Optical Wireless Data Center Networks," in *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2245-2259, March 2019.

[14] B. C. Chatterjee, N. Sarma and P. P. Sahu, "Priority based routing and wavelength assignment with traffic grooming for optical networks," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 4, no. 6, pp. 480-489, June 2012.

[15] Li, S., Zukerman, M., Wang, M., & Wong, E.W. (2015). Improving throughput and effective utilization in OBS networks. *Opt. Switch. Netw., 18*, 222-234.

[16] Y. Lim and J. E. Kobza, "Analysis of a delay-dependent priority discipline in an integrated multiclass traffic fast packet switch," in *IEEE Transactions on Communications*, vol. 38, no. 5, pp. 659-665, May 1990.