Evaluation of Optical Packet Switch as Edge Device Using OPNET Modeler

Amr Mohamed, Ayman Kaheel, Tamer Khattab, and Hussein Alnuweiri

Department of Electrical & Computer Engineering
University of British Columbia
Vancouver, B.C., CANADA V6T 1Z4
Email: amrm@ece.ubc.ca

Abstract

In this paper we propose a network architecture that enables the transfer of data over WDM optical networks transparently without the need of buffers in network core. The network architecture consists of optical packet switches (OPS) acting as edge devices and optical cross-connects (OXC) in the core. Furthermore, we introduce an aggregation capable OPS (AC-OPS) edge device and study its effect on network utilization and end-to-end delay.

To evaluate the performance of the proposed network architecture, we have built a sample network model using the OPNET Modeler, which provides a convenient simulation platform for WDM networks. Our simulation results show that the AC-OPS enhances the network utilization over the classical OPS. In addition, we show that the performance of the network architecture under study depends heavily on two proposed aggregation parameters, namely the maximum burst size and aggregation timeout.

1. Introduction

Recent advances in the last decade promote WDM networks to become the strongest candidate for providing faster networking infrastructure. The explosive growth of bandwidth for optical fibers will serve as the base for providing guaranteed QoS features that the Internet strives to fulfill. On the other hand, technological limitations in optical logic circuits and optical buffering devices slow down the evolution of optical switching techniques, and the establishment of differentiated optical services [2].

To allow for utilizing the huge fiber bandwidth while overcoming the technological limitations, the switching of optical packets inside the optical transport network (OTN) must be fast and must exploit the wavelength domain in generating multiple light-paths over the fiber line. For more efficient utilization, the network edge device has to be capable of aggregating the network traffic before feeding to the OTN.

Numerous switching techniques have been developed and several prototype architectures have been designed to fulfill these requirements. To achieve the high switching inside the OTN, packet-processing time must be minimum. One of the switching techniques that makes this possible is the optical cross connect (also called wavelength router). OXC switches represent traditional circuit switching in optical networks in a slightly more flexible manner. OXC switches are used to route the optical packets based solely on their optical and spatial characteristics. OXC differentiates data flows based on the input port (spatial) and the wavelength (optical) of the incoming optical packets [3].

Another switching technique which routes the optical packets based on their final destination is the optical packet switching (OPS). This is may be desired for better utilization and the more dynamic multiplexing of data. However, processing the optical packet to recognize the address will require electrical to optical conversion due to the technological limitation in the optical processing units. To minimize the delay imposed by the optical to electrical conversion and vice versa, the payload is separated from the header section for each packet so that the payload always remains in the optical domain while the header is converted to electrical for processing in each node and then back to optical to join the payload.

Optical burst switching (OBS) [1] is a variation of Optical packet switching. The basic idea of OBS is that a burst (collection) of data packets moving from the same source to the same destination is transmitted together over a path that is reserved, prior to the transmission, using a single control header. The path is then disconnected after the burst is sent. There are several variations of OBS techniques. The two main techniques that are used in OBS are Just In Time (JIT) OBS and Just Enough Time (JET) OBS. In both techniques the header is separated from the data and is sent over a dedicated separate channel. In JIT OBS, a sender requesting a path sends the control packet over a separate channel. Upon receiving the path request header, a switch allocates one of the available wavelengths at a certain output port. If there is no wavelength available, the request is rejected and the data packet is discarded. Meanwhile, after a certain delay T the source starts transmitting the data portion of the burst regardless of the status of the request. This technique belongs to a class of switching protocols called Tell and Go (TAG) switching [4][5]. In JET OBS, the reservation request will contain the information of when the path should be torn down. This allows for more efficient use of the bandwidth. If the node receiving the reservation request did not find a wavelength free to fulfill the request, it can lookup a wavelength that will become free in a reasonable time and schedule the burst transmission over this wavelength after it becomes free.

In this paper we propose new optical network architecture and explain the design and implementation of the simulation model that realizes this architecture. We also present the results of our performance analysis based on the network utilization and end-to-end (ETE) delay as criteria for measuring system performance. Section 2 explains the top-level network architecture and illustrates the data aggregation concept. Section 3 focuses on describing the network model components including OXC, OPS and the aggregation models. Section 4 presents our simulation results.

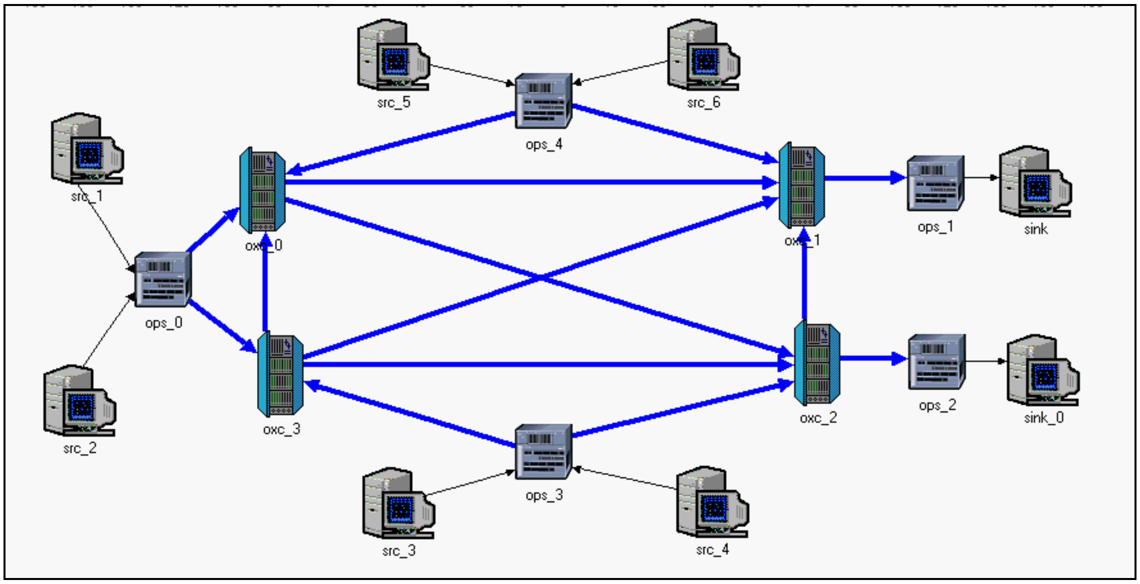

Figure 1: Optical network architecture

2. Optical network architecture

The proposed network architecture is shown in Figure 1. The architecture consists of a fully connected mesh of optical cross connect (OXC) in the core optical network and OPS at the optical network edge. As explained before, the OXC switches optical packets based solely on their wavelength. OXC uses pre-configured forwarding tables for switching optical packets from one input port to an output port. Hence, it creates a cross connection between a certain wavelength on an input port and an output wavelength on an output port. The group of these cross connections between different links will form the light path that the optical packets will travel through from the source to the destination as shown in **Figure 2**. While OXC provides fast and high switching capacity because it does not process incoming optical packet headers, it cannot efficiently utilize the optical network links alone because it only establishes multiple virtual circuits and does not adapt to network loads or optical links congestion.

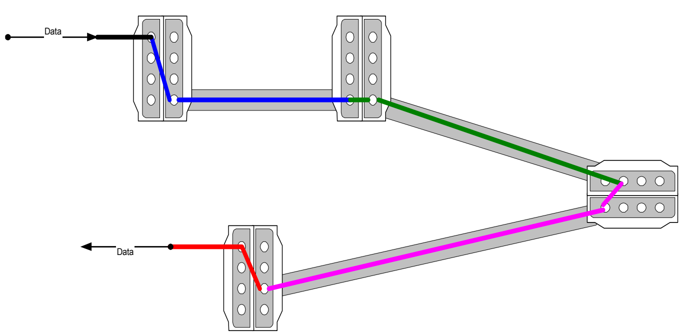

Figure 2: Light-path establishment in OXC switching

OPS switches, on the other hand, process the incoming optical packet headers to provide more bandwidth granularity. Then, it forwards packets based on the destination address. For optical packets forwarding, OPS uses lookup tables, which may be built dynamically using routing protocols. Therefore, OPS is capable of responding to network conditions and hence, provides more efficient utilization for optical links.

2.1. Aggregation in edge OPS

OPS as an edge device can also be used for more efficient link utilization. This is achieved by aggregating the data traffic from the end systems based on the final destination that the data is targeted for. OPS then switches the entire data burst to the OTN after the electrical to optical conversion. This scheme significantly minimizes the switching time required per packet and increases the whole network utilization especially in network overload conditions.

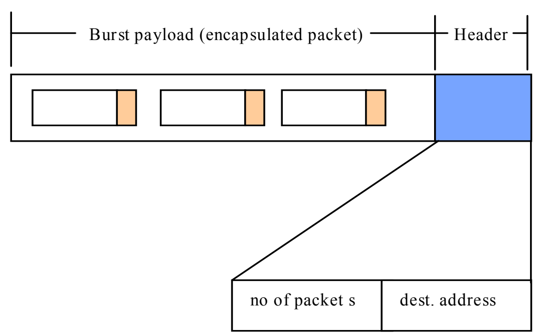

Figure 3: Data burst

OPS switches with aggregation are called aggregation capable OPS switches or (AC-OPS). The aggregation scheme is depicted in Figure 3. As the data packets arrive to the OPS input port, the OPS control unit concatenates them into one burst until the maximum burst size is reached. The entire burst is then converted to optical and fed to the OTN using one the OBS techniques explained before. For low packet rates, however, the burst may take long to be created, which imposes large delays to the data packets. This delay is called aggregation delay. To avoid large aggregation delay, an aggregation timeout may also be used during the aggregation process to stop forming the burst.

3. Network model

The sample network topology used for our simulation experiments is depicted in Figure 1. The network consists of a core part and external part. The network core, represents the OTN, consists of OXCs connected in a full mesh. OPSs with aggregation capabilities are used as the network edge devices. The OPS in this network works as a converter from electrical to optical domain at the network ingress, and works as a converter from optical to electrical domain at the network egress. Links connecting sources to ingress OPSs and those connecting egress OPSs to destinations carry one data channel with a capacity of 2.3 Gbps (OC-48). OTN internal links carry four data channels each with a capacity of 2.3 Gbps. We are assuming that all channels have zero bit error rate and propagation delays will be ignored. Although the network core physical topology is a fully connected mesh, the switching tables decide the logical topology. The logical topology of our network is depicted in Figure 4. The logical topology is chosen such that a reasonable load is created on some links without the need to increase source rates.

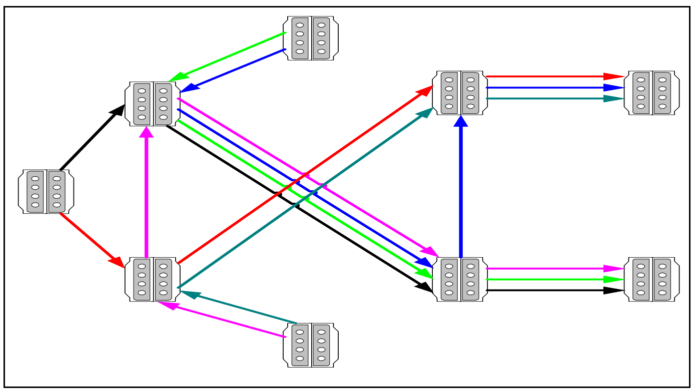

Figure 4: network logical topology

3.1. OXC model

The OXC node model is depicted in Figure 5. The switch model consists of four input/output ports each of which is capable of carrying up to four different wavelengths represented by packet streams in the model. The switch control unit is represented by the OXC processor. The control unit is responsible for reading the switching table, performing the wavelength conversion and relaying the optical packet to the output port according to the switching table. An example of the switching table is shown in Table 1. The main objective of the OXC model design is to keep the switching operation as simple as possible so that the switching delay can be ignored in the calculations and to be able to switch the packets transparent to its contents, which is a main requirement for core network switching as explained before.

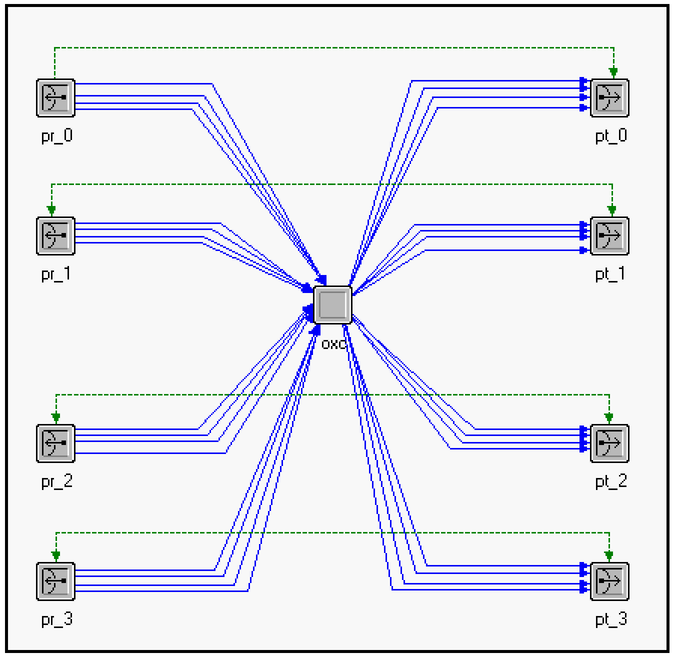

Figure 5: OXC node model

Input		Output	
Port	Wavelength	Port	Wavelength
0	0	1	0
0	1	2	0
0	2	3	0
1	0	0	0
1	1	2	1
1	2	3	1
2	0	0	1
2	1	1	1
2	2	3	2
3	0	0	2
3	1	1	2
3	2	2	2

Table 1: OXC forwarding table

3.2. OPS model

The node model of the edge OPS is depicted in Figure 6. The main role of the OPS is to forward the optical packets based on the destination address, which is one field in the packet header. Working as an edge device, the OPS acts as the interface between the OTN and the terminal workstations. Thus, the model must support two interfaces for input and output on the OTN side as well as the terminal workstations side in addition to the optical to electrical and electrical to optical conversion.

One of the critical issues that the OPS switch addresses is the packet contention or (external blocking). This happens when two or more optical packets are resolved to go to the same output port at the same time. This problem is partially resolved in optical networks because of the wide range wavelengths that can be used on one fiber line. However, the problem persists if all the wavelengths on the fiber are utilized and more than one packet using the same wavelength go to the same output port at the same time. The most common solution for this problem is to buffer the optical packets and send one at a time on the output port. This is done in optical switches by using fiber delay lines (FDLs) at each output port. FDLs are long fiber lines that delay the optical packets for a time period proportional to the FDL length. So, a group of FDLs with variable length can act as a FIFO queue.

Similar to OXC model, the OPS model consists of four input ports and four output ports represented by the point-to-point receivers and transmitters respectively. The four input ports include two ports serve as OTN interface, and they are called 'ingress' input ports. The other two input ports are interface to the terminal workstations and they are called 'egress' input ports. Each ingress port is connected with four channels represented by packet streams as shown in Figure 6. Each channel depicts one wavelength in the optical domain.

The OPS processor represents the switch control unit, which controls the input/output ports and performs the physical packet forwarding. The OPS processor loads two tables that describe the configured paths as shown in Figure 4. The two tables include the forwarding table and the switching table. Table 2 shows the global forwarding table for Egress OPSs. The forwarding table is used by the OPS to lookup the final destination for all the workstations connected on the OPS. This happens when the OPS is used as an egress forwarding packets to final destinations directly connect to the OPS. The switching table is used to forward the optical packets on a particular output stream on an output port as shown in Table 3. This happens when the OPS is used as ingress forwarding packets to OTN.

Some packet streams are connected with queues for resolving external blocking. Queues at the ingress output ports are also used for aggregation as explained in the next section. The queues on the ingress output ports are used for de-aggregation. This model simply incorporates the burst electrical to optical conversion, reads the burst header and split the burst into packets, which get forwarded to the OPS control unit for switching. The queues on the egress output ports are standard FIFO queues to resolve packet contention and model the de-queuing of the de-aggregated packets.

Egress address	Dest. address	Dest. port
1	0	2
2	1	2

Table 2: OPS forwarding table

Dest. address	Dest. egress	Dest. port	Dest. wavelength
0	1	0	0
1	2	1	0

Table 3: OPS switching table

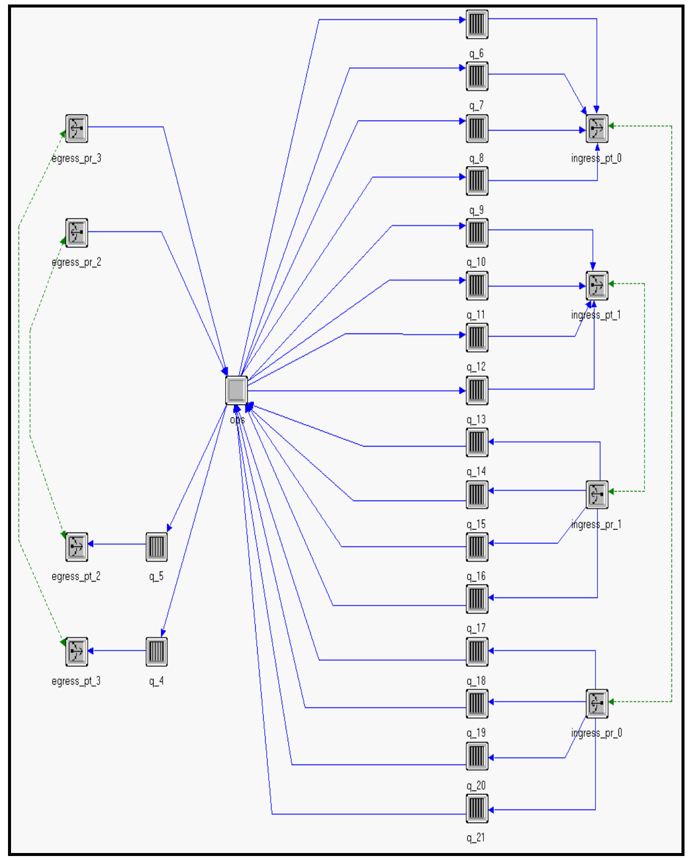

Figure 6: OPS node model

3.3. Aggregation model

In principle, the aggregation model's role is to assemble arriving packets into bursts, and send these bursts through the OTN. Two parameters control this assembling operation, namely maximum burst size and aggregation timeout. The maximum burst size controls the maximum number of packets in the burst. If the incoming traffic is bursty enough, then the maximum burst size will be reached in a reasonable time. However, if the traffic has a low arrival rate, the burst will have to wait till the maximum number of packets is reached. In order to avoid large aggregation delays, we use the aggregation timeout parameter. The aggregation timeout limits the maximum assembly delay a burst can encounter. In our implementation, both the maximum burst size and aggregation timeout are simulation parameters, in order to study their effects.

The aggregation process model is depicted in Figure 7. The aggregate state processes the packet arrival events. The state reads the egress number and classifies the packet to the proper burst. If there is no burst created for this egress, the state creates new burst and assigns the egress number to it. If the maximum burst size is reached or the timeout period has elapsed, the burst is marked as ready to be sent and a new burst is created for subsequent packets for this egress. If the destination wavelength at the destination port is free, the burst is converted to optical domain and sent after attaching the burst header. Otherwise, the burst is queued in electrical domain until the wavelength becomes free. The burst header consists of two fields, number of packets encapsulated in the burst and the address of the destination egress node as shown in Figure 3.

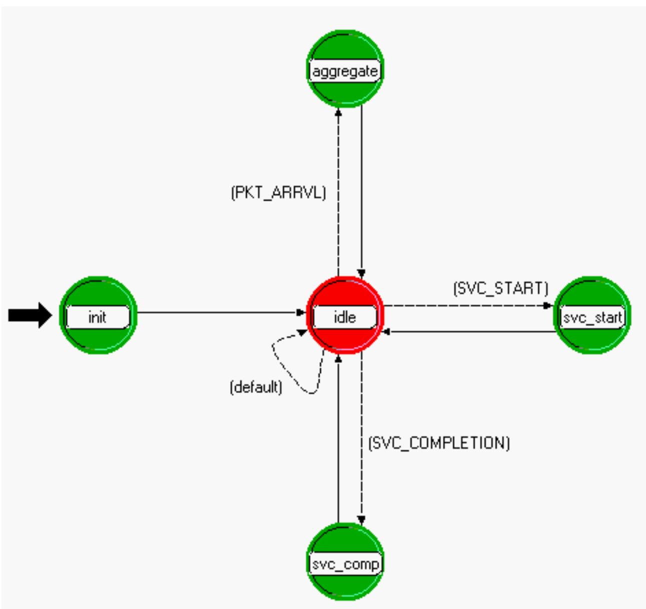

Figure 7: Aggregation process model

4. Simulation Experiments and results

In these simulation experiments, we study the effect of maximum burst size and timeout parameters on network utilization and end-to-end (ETE) delay. We also compare the performance of an aggregation capable optical packet switch with an optical packet switch that does not perform aggregation. Results obtained in this paper are supported by a mean-value mathematical analysis for the ETE delay and network utilization, which is reported in [6].

In the following, we denote the maximum burst size B_{max} and the aggregation timeout as *T*. All sources used in these experiments are exponential sources with mean *r*. We set the packet size to fixed size equal to 1024 bits/packet.

4.1. Effect of Maximum Burst Size

In this experiment we study the effect of varying the maximum burst size on both ETE delay and network utilization. We vary the maximum burst size and evaluate ETE delay and network utilization. For this experiment, we set *T= 1 sec* to neutralize the effect of aggregation timeout. We also choose r= 2 μ*sec*. Figure 8 and Figure 9 show the experiment results. Figure 8 shows that as B_{max} increases the network utilization increases until the utilization becomes almost constant after B_{max} reaches 51200 bits (50 packets).

However, Figure 9 shows that ETE delay continuously and linearly increases with the increase of B_{max}. This result is intuitive since the increase in B_{max} cause packets to be queued in the aggregation queue for longer period to wait for more packets to come until the burst size reach B_{max} and the larger B_{max} the longer the wait. Accordingly, it is important to find an operating area for B_{max} such that utilization is optimized while ETE delay is acceptable based on a specific application.

4.2. Effect of Aggregation Timeout

The effect of aggregation timeout is studied in this experiment. We change the aggregation timeout value and record the value of the network utilization and ETE delay. In this experiment we set B_{max} = 20480 (20 packets) and r= 2 μ*sec*. We choose the value of B_{max} such that it will not affect the experiment when *T* has small values. However it starts to take effect when *T* has relatively large values. Results are shown in Figure 10 and Figure 11.

Figure 10 shows that as the aggregation timeout increases the network utilization increases until *T* is equal to 40 μ*sec,* and beyond this point *T* has no effect on the network utilization. This can be understood if we observe that when *T* reaches 40 μ*sec,* it allows the maximum burst size to be reached. At which point B_{max} becomes the controlling factor. The same result can be seen from Figure 11. As the aggregation timeout increase the ETE increases, this is because a larger *T* allows for a larger aggregation delay and when *T* reaches 40 μ*sec,* the ETE becomes constant as the role of B_{max} becomes in effect.

4.3. Performance Comparison

In this experiment, we compare the performance of OPS with that of the Aggregation capable OPS in terms of ETE delay and network utilization. In the first part of this experiment, we change r and record ETE delay and utilization using different values of B_{max}. Results of this part are shown in Figure 12 and Figure 13. Figure 12 shows network utilization for different values of burst sizes including the case when aggregation is not used (Maximum burst size =1024 or 1 packet). The figure shows that the utilization is directionally proportional with 1/r. The OPS with aggregation capability performs much better in terms of utilization, and the larger the B_{max} used, the better utilization we can get. This is expected as we have shown in the previous experiment.

Figure 13 shows that the ETE delay is proportional to B_{max}/r for all burst sizes greater than 1024 (1 packet). For burst size equal to 1024, ETE delay has a minimum constant value for all values of inter-packet arrivals, which is logical because the packet does not encounter any aggregation delay in this case. For higher burst sizes, the aggregation delay increases proportionally with 1/*r* causing the ETE delay to increase.

Part two of this experiment focuses on the effect of time-out on the aggregation capable OPS. The value of B_{max} is set to 20480 bits (20 packets). Results are shown in Figure 14 and Figure 15. Figure 14 shows utilization vs. *r* for three values of *T*. In case of *T* = 1e-5, we can see that the utilization deteriorates with the increase of *r*. This is due to the fact that as *r* increases, the burst size decreases, which in turn decreases the utilization. For the other two values of the aggregation timeout *T*, the utilization was not severely affected by the increase in *r*. This is because both values of *T* (50 and 100 µ*sec*) are large enough to include a large number of packets for all values of *r*. For example, when *r* = 10 µ*sec*, the burst will consist of 5 and 10 packets for *T* = 50 and 100 µ*sec* respectively.

In order to understand the results in Figure 15, we should keep in mind that the average packet ETE delay is the sum of three terms, namely: average aggregation queuing delay + burst transmission delay + de-queuing delay (at the egress OPS). Consider the case of *T*=50 µ*sec*. For low values of 1/*r*, the burst size becomes the dominant factor, as the number of packets in the burst is constant and equal to 20 packets. However the aggregation queuing delay is affected by the arrival rate *r* causing the ETE delay to increase as 1/*r* increases. When 1/*r* is close to 2.5, the effect of time-out starts to play a role when the 20 packets burst takes around 50 µ*sec*. After that the number of packets in the burst starts to decrease and cause the transmission delay to decrease. This explains the small decrease in the ETE delay shown on the figure. As 1/*r* increases the aggregation delay increases, the transmission delay and the de-queuing delays decrease. Because the aggregation delay is directly proportional to 1/*r*, it becomes more dominant and causes the ETE delay to increase. For 1/*r* more that 50 µ*sec*, the burst size is always one packet causing the aggregation delay to be 50 µ*sec* and the transmission delay is constant. Because the de-queuing delay is dependent on the number of sources sending on the network link (3 sources, see Figure 1), the ETE delay takes only slight changes until it saturates at around 50 µ*sec* when 1/*r* is high enough for the de-queuing delay to be constant.

5. Conclusion and future work

In this paper we presented a scalable and modular framework for simulating optical networks, which is designed to be expandable into further scenarios and protocols for optical networks. We have used the simulation model to evaluate AC-OPS as one of the state of the art optical switching techniques. We have studied effect of aggregation on both end-to-end delay and network utilization.

Our simulation results show that AC-OPS promises better utilization for the network. Of course this doesn't come for free, as can be seen from the increase in the end-to-end delay. However, the simulation results show that the network operator can carefully choose the aggregation parameters such that the utilization can be increased without significantly increasing the end-to-end delay. This can be done by avoiding the points where the delay reaches its maximum values for a given utilization value.

This work can be extended in the future to include QoS critical algorithms and see how these algorithms can be tuned or modified to enhance the performance of optical networks. A further extension for this work is to add the signaling techniques required for OBS to work at its maximum dynamic extents. A third dimension of extending this work is to exploit different switching architectures such as broadcast-and-select and see how these architectures can be modified to enhance the overall performance of the network

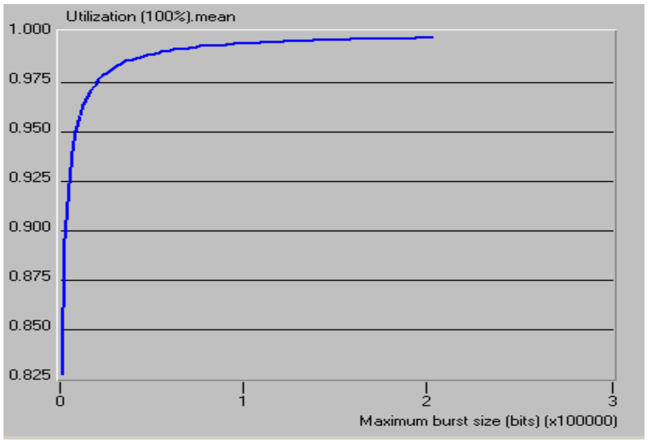

Figure 8: Effect of *B* on utilization

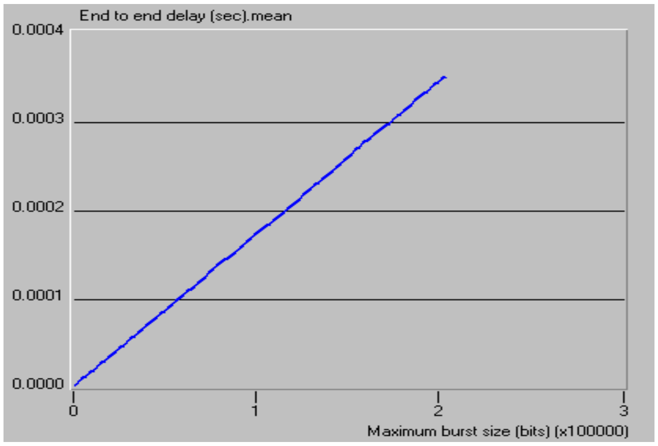

Figure 9: Effect of *B* on ETE delay

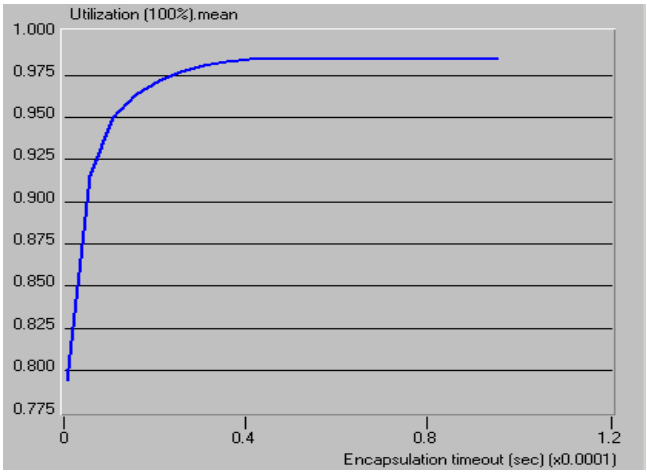

Figure 10: Effect of *T* on utilization

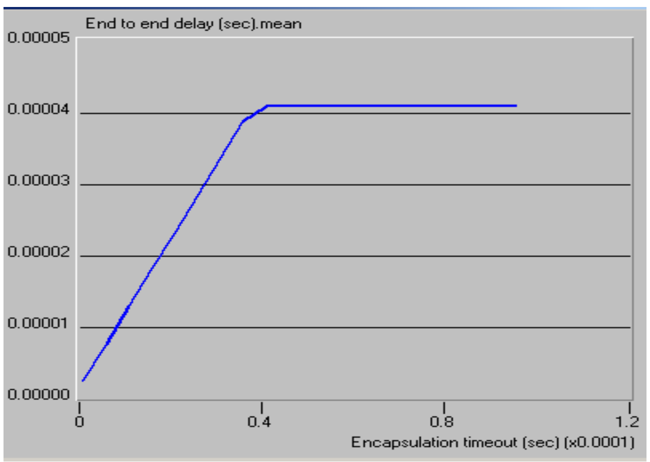

Figure 11: Effect of *T* on delay

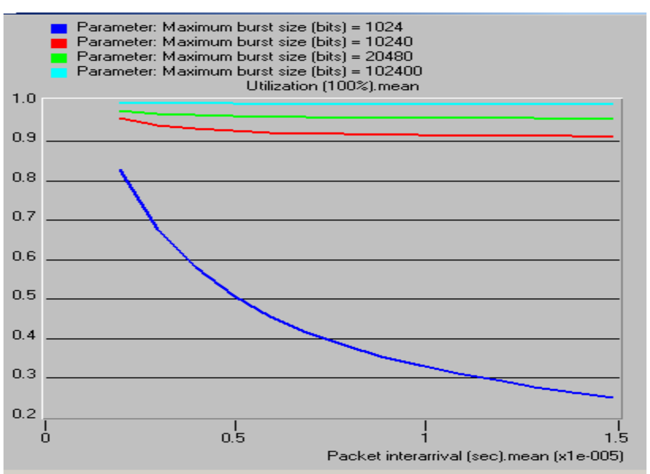

Figure 12: Utilization vs. inter-arrival time for different *B*

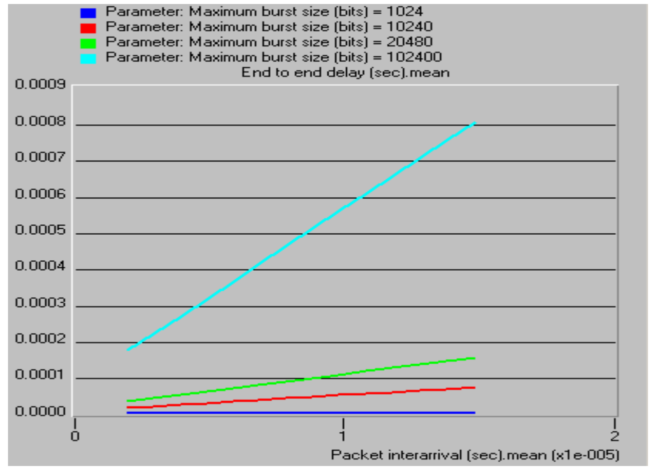

Figure 13: ETE Delay vs. inter-arrival time for different *B*

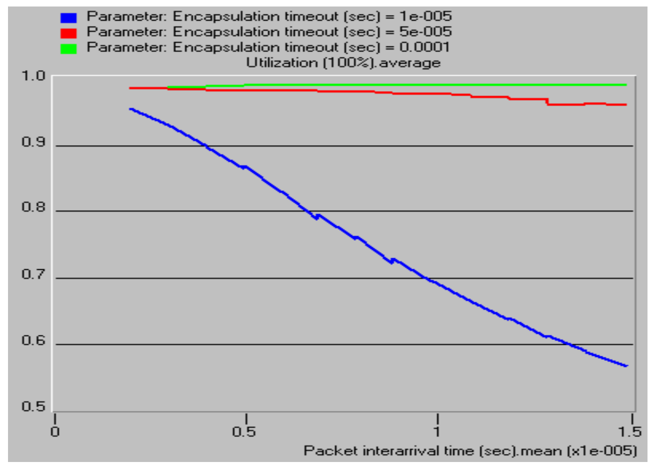

Figure 14: Utilization vs. inter-arrival time for different *T*

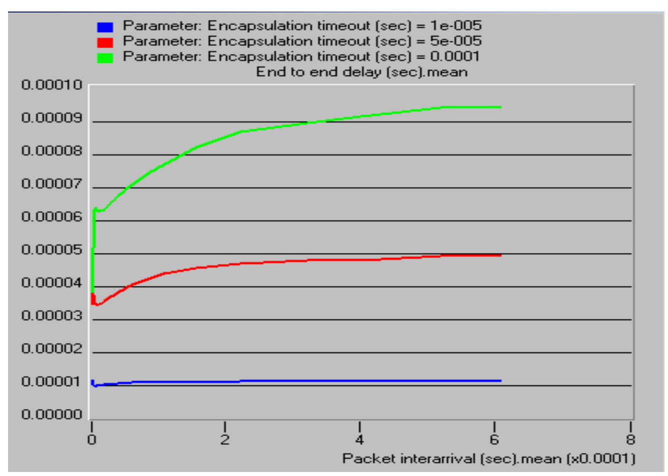

Figure 15: ETE Delay vs. inter-arrival time for different *T*

References

[1] M. Yoo and C. Qiao, “A new optical burst switching protocol for supporting quality of service,” in SPIE Proceedings, All Optical Networking: Architecture, Control and Management Issue, vol. 3531, pp. 396–405, November 1998.

[2] N. Golmie, T. D. Ndousse, and D.H. Su, “A differentiated optical services model for WDM networks,” IEEE Communication Magazine, pp. 68–73, February 2000.

[3] Jaafar M.H. and Hussein T. Muftah, “Technologies and Architectures for Scalable Dynamic Dense WDM Networks”, IEEE Comm. Mag. Feb. 2000.

[4] Klaus Dolzer et al, Evaluation of Reservation Mechanisms for Optical Burst Switching, International Journal of Electronics and Communications (AEÜ), Vol 55 No.1, Jan 2001.

[5] Ilia Baldine et al, JumpStart: A Just-in-Time Signaling Architecture for WDM Burst-Switched Networks, IEEE Comm. Mag. Feb. 2002

[6] T. Khattab, A. Mohamed, A. Kaheel, “Optical Packet Switching with Packet Aggregation”, submitted to SOFTCOM 2002.