Survey on Optical Burst Switching in WDM Networks

Pushpendra Kumar Chandra, Ashok Kumar Turuk, Bibhudatta Sahoo
National Institute of Technology Rourkela, Orissa, India
chandrap@nitrkl.ac.in, akturuk@nitrkl.ac.in, bdsahu@nitrkl.ac.in

Abstract—Optical burst switching (OBS) is the promising area it can meet the bandwidth requirement of bandwidth intensive application. OBS is switching technique for the next generation optical networks. However, there are certain issues such as burst aggregation, scheduling, contention resolution and QoS that needs to be addressed in OBS. In this paper, we summarize the results of current research trends on optical burst switching networks. We present the classification of different scheduling algorithms and contention resolution techniques.

I. INTRODUCTION

There have been a phenomenal increase in the demand for bandwidth over the years due to rapid growth in the number of Internet users and increase in bandwidth intensive applications such as voice-over-IP, video conferencing, interactive video on demand, and many other multimedia applications [1]. To meet the ever growing demand of bandwidth, copper cables were replaced by optical fibers in both the access networks as well as in the backbone networks [2]. Optical fiber not only supports huge bandwidth but also have other advantages too such as lower bit-error rate, no interference problem and security advantage [3]. In first generation optical networks, optical fibers provided only point-to-point connections. Entire potential of the fiber could not be utilized, because the electronic routers operate at a much lower speed than the fiber capacity. Wavelength division multiplexing (WDM) technology, were deployed in the second generation optical networks. WDM divides the available bandwidth of the fiber into number of non-overlapping wavelength channels each operating at electronic speed. To carry IP traffic over WDM networks three switching technologies have been studied: optical circuit, packet switching and burst switching. Optical circuit switching and packet switching have their own limitations when applied to WDM networks [4]. Circuit switching is not bandwidth efficient unless the duration of transmission is greater than the circuit establishment period. It is shown that establishment of circuits (lightpaths) in optical networks is an NP-hard problem [5-7]. On the other hand packet switching is hop-by-hop store and forward scheme and needs buffering and processing at each intermediate node [8]. It is flexible and bandwidth efficient. However, technology for buffering and processing in optical domain is yet to mature for this scheme to be commercialized [9, 10]. Fiber delay lines (FDL) have been proposed in literature to provide buffering. However, FDL have limited buffering capability and support only for a fixed duration [11].

In this context optical burst switching (OBS) [12-16] is emerging as the alternative switching techniques, which combines the advantages of both circuit switching and packet switching. In OBS, a burst is the basic switching entity. Burst is a variable length data packet, assembled at an ingress router by aggregating a number of IP packets, which may be received from a single host or from multiple hosts belonging to the same or different access networks. A burst has two components: control and payload [17, 18]. The control packet carries the header information. Thus, the control component incurs an overhead, referred to as control overhead. Payload is the actual data transmitted. Functional diagram of OBS [19] is shown in Fig. 1. The ingress node is responsible for burst assembly, routing, wavelength assignment, scheduling of burst at edge node. The core node is responsible for signaling, scheduling, resolving contention. The egress edge node is responsible for disassembling the burst and forwarding the packet to higher network layer.

Fig. 1. OBS Functional Diagram

Comparison of three switching technology is given in Table I [20].

<table>
<thead>
<tr>
<th>Switching</th>
<th>Bandwidth Utilization</th>
<th>Latency</th>
<th>Optical Buffering</th>
<th>Overhead</th>
<th>Adaptively</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuit</td>
<td>Low</td>
<td>High</td>
<td>Not Required</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Packet</td>
<td>High</td>
<td>Low</td>
<td>Required</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Optical</td>
<td>High</td>
<td>Required</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

II. ARCHITECTURE OF OBS

An optical burst-switched network composed with OBS nodes that are interconnected via optical fiber link. Architecture of an OBS networks is shown in Fig. 2: OBS network having two types of nodes: edge node and core node [2, 21, 22]. Edge nodes are at the interface between the electronic and optical domain. Edge nodes can be an ingress or egress node. Packets are assembled into bursts at ingress edge node, which are then routed through the OBS network and disassembled back into packets at egress edge node. A
core node is mainly composed of an optical switching matrix and a switch control unit which is responsible to forward data burst.

A node in OBS network consists of both optical and electronic components. The optical components are multiplexers, demultiplexers and an optical switching network. The electronic components are input modules (IM), output module (OM), a control burst router (CBRT), and a scheduler [21]. An optical burst switch control unit transfers an incoming data burst from an input port to its destination output port.

When an edge node intends to transmit a data burst, it sends a control packet on the control wavelength to a core node. At core node, the control packet on the control channel is input to the corresponding IM, which converts the control packet into electronic form. The control fields are extracted from the control packet. The CBRT uses these control fields to determine the next outgoing fiber for the corresponding payload by consulting a routing table maintained locally. The control packet is scheduled for transmission onto the selected outgoing link by the scheduler and the control packet is buffered until the scheduled time. The scheduler maintains a control packet queue. The scheduler also reserves wavelength on the determined links for the upcoming payload. The control packet is then forwarded on the OM, which updates its control fields and transmits it to the selected outgoing fiber using the optical transmitter. Just before the payload arrives, the switching element in the node is configured to connect the input port to the corresponding output port for the entire duration of the burst transmission.

If the control packet is unable to reserve the wavelength then the control packet as well as payload is dropped.

III. BURST ASSEMBLY SCHEMES

Burst assembly is the process of assembling all incoming data burst at the edge node. Most widely used burst assembly schemes are threshold based and timer-based [23, 24]. In a timer based scheme, a timer is started at the initialization of burst assembly. A data burst containing all packets in the buffer are generated when the timer exceeds the burst assembly period. A large time-out value results in a large packet and higher buffering delay at the edge node. On the other hand, a too small value results in too many small bursts and a high electronic processing load.

In a threshold-based scheme, a burst is created and sent into the OBS network when the total size of the packets in the queue reaches a threshold value. The shortcoming of the threshold-based scheme is that it does not provide any guarantee on the assembly delay that packets will experience. The choice of burst assembly algorithms depends on the type of traffic being transmitted. Timer-based algorithms are suitable for time-constrained traffic such as real-time applications because the upper bound of the burst assembly delay is limited. For a time-insensitive application such as file transmission, to reduce the overhead of control packets and increase OBS transmission efficiency, a threshold based scheme may be more appropriate.

IV. WAVELENGTH RESERVATION SCHEMES

Wavelength reservation refers to when and how the bandwidth is reserved and release. The reservation schemes in OBS network is adopted from ATM Block Transfer (ABT) [25]. There are two versions of ABT: ABT with delayed transmission and ABT with immediate transmission.

In an immediate transmission reservation scheme, an output wavelength is reserved for a payload immediately after the arrival of the corresponding control packet; if a wavelength cannot be reserved at that time, then the setup message is rejected and the corresponding data burst is dropped [16]. In a delayed reservation scheme, the control packet and the payload are separated in time by an offset value in order to accommodate the processing of the control packet. An output wavelength is reserved for a data burst just before the arrival of the first bit of the data burst. If, upon arrival of the setup message, it is determined that no wavelength can be reserved at the appropriate time, then the setup message is rejected and the corresponding data burst is dropped [16]. These two techniques have been adopted in OBS. Depending on bandwidth reservation, offset time and control management, three schemes for OBS implementation have been proposed: Tell-and-go (TAG) [16], Just-in-time (JIT) [15,26] and Just-enough-time (JET) [27].

A. Tell-And-Go(TAG)

This is an immediate reservation scheme. In TAG, the control packet is transmitted on a control channel followed by a payload, on a data channel with zero or negligible offset. The payload is buffered using fiber delay line (FDL) while the control packet is processing at each intermediate node. If wavelength reservation is successful then the payload is transmitted along the reserved channel else the data burst is dropped and a negative acknowledgment (NAK) is sent to the source. The source node sends a control packet after transmitting the payload to release the reserved resources along the path. The drawback of this scheme is availability of optical buffer. FDL can hold data only for a fixed duration and can not accommodate data burst of variable size. Furthermore, loss of control packet to release reserved resources result in wastage of bandwidth [21, 26].
B. Just-in-Time(JIT)

JIT is an immediate reservation scheme. Here, nodes reserve the resources as soon as the control packet is processed. Source transmits the payload after an offset time which is greater than the total processing time of control packet at intermediate nodes. If the resource is not available, the data burst is dropped. The difference between JIT and TAG is that in JIT the buffering of the payload at each node is eliminated by inserting a time slot between the control packet and the payload. The time slot is equal to the offset time. Since the bandwidth is reserved immediately after processing the control packet, the wavelength will be idle from the time the reservation is made till the first bit of the payload arrives at the node. This is because of the offset between the control packet and the payload. An in-band-terminator is placed at the end of each data burst, which is used by each node to release the reserved wavelength after transmitting the payload [6, 26].

C. Just-Enough-Time(JET)

JET is a delayed reservation scheme. Here, the size of the data burst is decided before the control packet is transmitted by the source. The offset between control packet and payload is also calculated based on the hop count between the source and destination. At each node, if bandwidth is available, the control packet reserves wavelength for the upcoming data burst for a fixed duration of time. The reservation is made from the time when the first bit of payload reaches the node till the last bit of payload is transmitted to the output port. This eliminates the wavelength idle time. This is the basic difference between JET and JIT. Since the wavelength is reserved for a fixed duration, there is no need for explicit release of reserved resources along the path. Since there is no wastage of bandwidth in this scheme, channel utilization is higher than other schemes.

TAG and JIT schemes are significantly simpler than JET since they do not involve complex scheduling or void filling algorithms. Previous studies have shown that JET performs better than either JIT or TAG in terms of burst loss probability [3, 16, 27-29].

V. BURST SCHEDULING ALGORITHM

When a control packet arrives at a core node, a wavelength channel scheduling algorithm is used to determine a wavelength channel on an outgoing link for the corresponding data burst. The information required by the scheduler such as the expected arrival time of the data burst and its duration are obtained from the control packet. The scheduler keeps track of the availability of time slots on every wavelength channel. It selects one among several idle channels. The selection of wavelength channel needs to be done in an efficient way so as to reduce the burst loss. At the same time, the scheduler must be simple and should not use any complex algorithm, because the routing nodes operate in a very high speed environment handling a large amount of burst traffic. A complex scheduling algorithm may lead to the early data burst arrival situation wherein the data burst arrives before its control packet is processed and eventually the data burst is dropped [3].

In this section we discuss various scheduling algorithms proposed in literature [30,31]. These algorithms differ in their complexity and performance in terms of burst loss. A wavelength channel is said to be unscheduled at time t when no data burst is using the channel at or after time t. Algorithms which consider unscheduled channels are called Horizon algorithm. A channel is said to be unused for the duration of voids between two successive data bursts and after the last data burst assigned to the channel. Algorithms which consider voids within channels are called void filling algorithm. According to scheduling strategy used scheduling algorithms can be classified as follows:

- Horizon or without void filling [30].
- With void filling [31].

Representative of Horizon algorithms are: First Fit Unscheduled Channel (FFUC) [30-33], Latest Available Unused Channel (LAUC) [6, 33] and that of void filling algorithms are: First Fit Unscheduled Channel with Void Filling (FFUC-VF) [31], Latest Available Unused Channel with Void Filling (LAUC-VF) [32, 34, 35] and Minimum End Void (Min-EV) [35]. Working of algorithms is illustrated with the help of Fig. 3. In Fig. 3: control packet arrives at a node at time \(t_{cb}\). Duration of payload is \(t_{burst}\) and the offset time for the data burst is \(t_{offset}\). The offset time is calculated as:

\[
\text{offset} = H \times \text{t}_{\text{up}}
\]

where \(H\) is number of hops from source to destination and \(\text{t}_{\text{up}}\) is the time required for processing and switching the control packet. The time at which the first bit of payload arrives at the node is \(t_{cb} + t_{offset}\) and the last bit arrive at \(t_{cb} + t_{offset} + t_{burst}\). We define unscheduled channel and void channel as following: unscheduled channel: A wavelength channel is said to be unscheduled at time t when no data burst is using the channel at or after t. void channel: If a channel is unused for a duration between two successive data bursts.

![Fig. 3. Illustration of Burst Scheduling Algorithms](image)

A. First Fit Unscheduled Channel (FFUC)

First fit unscheduled channel (FFUC), selects an unscheduled channel for an incoming data burst [31, 32, 36]. FFUC, keeps the unscheduled time for each data channel.
When a control packet arrives, the FFUC algorithm searches all data channels in a fixed order and assigns the data burst to the first channel that is available at or after the arrival time of the payload. In Fig. 3: when a control packet arrives at a time $t_{cb}$, the scheduling algorithm searches for all unused channels. Available unscheduled channels are channel 1 and channel 2. FFUC selects channel 1, since this is the first available channel 1 and the channel 1 is reserved for the duration $T_{duration} = [t_{cb} + offset, t_{cb} + offset + burst]$ Advance of the algorithm is speed due to the relatively small number of channels that it checks. The best implementation of the FFUC scheduling algorithm takes $O(\log n)$ time to schedule a data burst, where $n$ is the number of data channels [21, 37]. Disadvantage of the algorithm is low network resource utilization due to following reasons: (1) does not consider voids that may appear between two already scheduled data bursts as a possible place for fitting the incoming data burst. (2) stops after first available channel.

B. Latest Available Unscheduled Channel (LAUC)

Latest available unscheduled channel (LAUC), selects an unscheduled data channel where the void created between consecutive scheduling of data bursts is minimum [31, 36]. In Fig. 3: channel 1 and 2 are two unscheduled channel at $t_{b}$. Scheduling on channel 1 creates a void ($t_{b} - t$) and in 2 is ($t_{b} - t$). Since ($t_{b} - t$) > ($t_{b} - t$), LAUC selects channel 2 for scheduling. LAUC has the same complexity as that of FFUC. In addition, LAUC utilizes the network resources better than FFUC.

C. First Fit Unscheduled Channel with Void Filling (FFUC-VF)

In First fit unscheduled channel with void filling (FFUC-VF) [36], all possible voids are found and the payload is scheduled on the first available void that is suitable for transmission. In Fig. 3: voids are available on the channel 3, 4, 5 and the duration of voids are ($t_{s} - t_{b}$), ($t_{s} - t_{b}$) and ($t_{s} - t_{b}$). FFUC-VF selects the channel 3 to schedule the data burst, because channel 3 is the first available void channel. If n is the number of data bursts currently scheduled on every data channel, then a binary search algorithm takes $log n$ time to check that the data channel is eligible or not. Thus the time complexity of the FFUC-VF algorithm is $O(w \log n)$, where $w$ is the number of data channels [21].

D. Latest Available Unscheduled Channel with Void Filling (LAUC-VF)

Latest available unscheduled channel with void filling (LAUC-VF) [34,36], searches all data channels to find an available void channel for the time interval ($t_{b} + offset$) and ($t_{b} + offset + burst$). Then select a channel, such that placement of new data burst create minimal void between newly arrival data burst start time and previous scheduled data burst end time. In Fig. 3: channel 3, 4, 5, 6 has such void. The difference between start time of newly arrival data burst and already scheduled data burst whose end time is prior to the start time of new data burst on the channels 3, 4, 5 and 6 are: ($t_{b} + offset - t_{b}$), ($t_{b} + offset - t_{b}$), ($t_{b} + offset - t_{b}$) and ($t_{b} + offset - t_{b}$) respectively. LAUC-VF select channel having minimum of the above time difference. So it selects channel 4 to schedule the incoming data burst. To implement LAUC-VF, switching control unit have to store usage information of all data channels. That makes LAUC-VF more complex compared to that of FFUC and LAUC. But it has higher network resource utilization.

E. Minimum End Void (Min-EV)

A variation of LAUC-VF algorithm is Minimum end void (Min-EV) [35]. It searches all data channels to find an available void channel to schedule the newly arrival data burst. Then, select a channel, such that placement of new data burst create minimal void between already scheduled data bursts start time and newly arrival data bursts end time. In Fig. 3: channel 3, 4, 5, 6 has such void. The difference between start time of already scheduled data burst and end time of newly arrival data burst on channel 3, 4, 5 and 6 are: ($t_{s} - (t_{b} + burst)$), ($t_{s} - (t_{b} + burst)$), ($t_{s} - (t_{b} + burst)$) and ($t_{b} - (t_{b} + burst)$) respectively. Min-EV selects a channel having minimum of the above value. Therefore, channel 5 is selected.

VI. CONTENTION RESOLUTION TECHNIQUES

Contention occurs when more than one data burst try to reserve the same wavelength channel on an outgoing link. In electronic network, contention is resolve by buffering the contending packets. In OBS network when contention occurs one of contending data burst is allowed to reserve the channel, for other data bursts one or a combination of the following contention resolution technique can be applied.

A. Optical Buffering

Optical buffering is achieved through the use of fiber delay lines (FDL). In optical network, fiber delay line (FDL) is currently the only way to implement optical buffering. To resolve contention using FDL, one of the contending data burst is passed through FDL. But it has several limitations. FDL are bulky and require over a kilometer of fiber to delay a single packet for 5 usec. [11], provide only a fixed delay [25] and data leave the FDL in the same order in which they entered [38]. Delay lines are commercially not viable due to the above drawbacks. In general, FDL can be used with other schemes to improve the performance.

B. Wavelength Conversion

Wavelength conversion is the process of converting a wavelength on an incoming channel to another wavelength on an outgoing channel [39-41]. To resolve contention using this method, a contending data bursts wavelength is shifted to another wavelength on the designated output link. Thus it increases wavelength re-usability. The following are the different categories of wavelength conversion:

- **Full conversion:** Any wavelength shifting is possible. Lightpaths can be converted regardless of their wavelength.
• **Limited conversion**: Wavelength shifting is restricted so that not all combinations of lightpaths may be connected.
• **Fixed conversion**: This is a restricted form of limited conversion such that, for each node, each lightpath may be connected to exactly one predetermined lightpath on all other links.
• **Sparse wavelength conversion**: In this scheme, networks are comprised of a mixture of nodes having full and no wavelength conversion capabilities.

C. Deflection Routing

Deflection routing is the approach of resolving contention by routing a contending burst to an output port other than the intended one. In WDM networks, where buffer capacity is limited and wavelength conversion is not feasible [18], the implementation of deflection routing may be necessary so as to reduce burst loss. Deflection routing is generally not preferred in electronic packet switched networks due to out of sequence delivery of packets but is desirable in OBS networks assuming that the offset time is sufficient. No extra hardware is required to perform deflection routing. In deflection routing, one of a contending data burst is sent to a different output port and then follows an alternative route to the destination [38, 40, 42].

D. Burst Segmentation

Burst segmentation is a promising solution to contention resolution. A data burst is composed of a number of segments. When contention occurs the burst is divided into two segments. One segment is dropped while the other is transmitted. The dropped segment can then be retransmitted.

Table II gives a comparison between four contention resolution techniques reported in [25].

<table>
<thead>
<tr>
<th>Contention Resolution</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength Conversion</td>
<td>The most efficient solution</td>
<td>Immature and expensive</td>
</tr>
<tr>
<td>FDL buffering</td>
<td>Simple</td>
<td>Increasing end-to-end delay</td>
</tr>
<tr>
<td>Deflection routing</td>
<td>No extra hardware requirement</td>
<td>Out of order arrival</td>
</tr>
<tr>
<td>Burst Segmentation</td>
<td>Lower packet loss ratio</td>
<td>Complicated control handling requirement</td>
</tr>
</tbody>
</table>

The policy for selecting which data bursts to drop is referred to as the soft contention resolution policy and is aimed at reducing the overall burst loss rate, and consequently enhancing link utilization. Several soft contention resolution algorithms have been proposed in [44], including the shortest-drop policy [45] and look-ahead contention resolution [46]. These contention resolution policies are considered as reactive approaches in the sense that they are invoked after contention has occurred. An alternative approach to reduce network contention is by proactively attempting to avoid network overload through traffic management policies [44].

VII. OBS Simulation Tools

Evaluating the OBS network performance, OBS simulators are mostly used by researchers. Four most relevant simulators used in scientific and academic research to model OBS networks are Network Simulator 2 (ns-2) [47] (i) OBS-ns Simulator [48] (ii) NCTUns (iii) OBSSimulator. Table III details the summary of simulator tools reported in [49].

<table>
<thead>
<tr>
<th>OBS Protocols</th>
<th>OBS ns</th>
<th>NCTUns</th>
<th>OBSsimulator</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBS Parameters</td>
<td>14</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>Network Traffic Generation</td>
<td>Application code</td>
<td>Real-life protocol stacks (TCP/IP)</td>
<td>Application code</td>
</tr>
<tr>
<td>Model Building</td>
<td>Script</td>
<td>Graphical model construction</td>
<td>Script</td>
</tr>
<tr>
<td>Output Analysis</td>
<td>Plot graphs</td>
<td>Plot graphs</td>
<td>Console output</td>
</tr>
<tr>
<td>Animation or Real-time Viewing</td>
<td>Animation</td>
<td>Animation</td>
<td>Not available</td>
</tr>
<tr>
<td>Support/Training</td>
<td>Available</td>
<td>Available</td>
<td>Not Available</td>
</tr>
<tr>
<td>Learning Curve</td>
<td>Steep</td>
<td>Short</td>
<td>Short</td>
</tr>
<tr>
<td>Installation Difficulty</td>
<td>Medium</td>
<td>Hard</td>
<td>Easy</td>
</tr>
<tr>
<td>Programming Language</td>
<td>C++</td>
<td>C++</td>
<td>C++</td>
</tr>
<tr>
<td>System Requirements</td>
<td>Linux</td>
<td>Linux</td>
<td>Linux</td>
</tr>
<tr>
<td>Download</td>
<td>Available</td>
<td>Available</td>
<td>Not Available</td>
</tr>
<tr>
<td>License Type</td>
<td>Free for research, commercially not available</td>
<td>Free for research, commercial licenses available</td>
<td>Private use</td>
</tr>
</tbody>
</table>

VIII. Conclusions

Optical burst switching is the new switching paradigm for the next generation optical network. In OBS control packets are decoupled from the data packets and are sent in different channels. Control information is sent on control channel and data packets are sent on data channels. Some of the research issues identified in OBS network are burst assembly and disassembly, burst scheduling, contention resolution and QoS. In this paper, we have tried to aggregate the work, reported by researchers in OBS networks.

ACKNOWLEDGMENT

This work is supported by the Department of Science and Technology (DST), under grant no. SR/S3/EECE/044/2007.

REFERENCES