A Layered Architecture for Supporting Optical Burst Switching

Farid Farahmand†, Vinod M. Vokkarane‡, Jason P. Jue‡

† ALCATEL Research and Innovation Center, Plano, Texas, U.S.A
‡ Advanced Networks Research Labs,
Department of Computer Science, The University of Texas at Dallas, Richardson, Texas
farid.farahmand@alcatel.com, {vinod, jjue}@utdallas.edu

Abstract: In this article, we define a new layered architecture for supporting optical burst switching in an optical core network. The architecture takes into account both the control plane as well as the data plane. We describe the functionality and the primary protocols that are required at each layer, and we explain how the layers interact with each other in the proposed architecture. Finally, we provide an illustrative example of an end-to-end data transmission using the layered architecture.

Keywords – IP, DWDM, Optical Burst Switching, Ingress Node, Egress Node.

1   Introduction

The amount of raw bandwidth available on fiber optic links has increased dramatically with advances in dense wavelength division multiplexing (DWDM) technology; however, existing optical network architectures are unable to fully utilize this bandwidth to support highly dynamic and bursty traffic. As the amount of bursty Internet traffic continues to grow, it will become increasingly critical to develop new architectures to provide the flexible and dynamic bandwidth allocation to support this traffic.
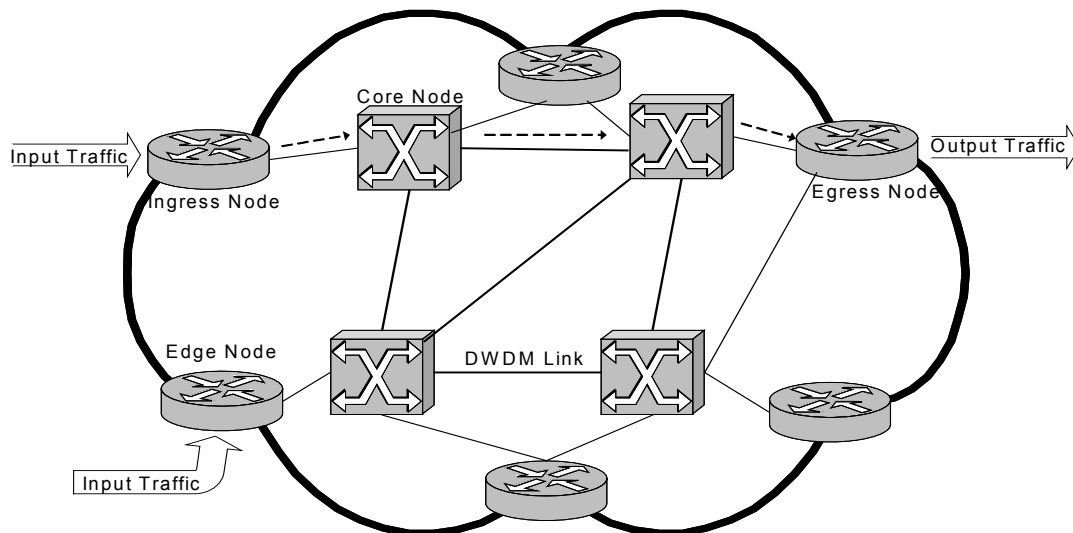
Figure 1. Optical transport network.

An optical transport network consists of a collection of edge and core nodes as shown in Fig. 1. The traffic from multiple client networks is accumulated at the ingress edge nodes and transmitted through high capacity DWDM links over the core. The egress edge nodes, upon receiving the data, provide the data to the corresponding client networks. The three prominent optical transport networks architectures proposed to carry traffic over the optical core are optical circuit switching (OCS) (or wavelength-routed networks), optical packet switching (OPS), and optical burst switching (OBS). These switching techniques primarily differ based on how resources are allocated in the core and the degree of granularity for the resource allocations.

In OCS networks, an all-optical connection, referred to as a lightpath [1], is established to create a logical circuit between two edge nodes across the optical core. These lightpaths may be established dynamically as connection requests arrive to the network, or they may be provisioned statically based on estimated traffic demands. While OCS is suitable for constant rate traffic such as voice traffic, it may be unsuitable for highly dynamic traffic. Furthermore, as lightpaths must be established using a two-way reservation scheme that incurs a round-trip delay, the high overhead of connection establishment may not be well-suited for short bursts of traffic. Also, under bursty traffic, sufficient bandwidth must be provisioned to support the peak traffic load, leading to inefficient network utilization at low or idle loads.

In OPS networks [2], data is transmitted in the form of optical packets which are transported across the optical core without conversion to electronics at intermediate core nodes. OPS can provide dynamic bandwidth allocation on a packet-by-packet basis. This dynamic allocation leads to a high degree of statistical multiplexing which enables the network to achieve a higher degree of utilization when the traffic is variable and bursty. However, there are many technical challenges to implementing a practical OPS system. One of the limitations of OPS networks is that it is difficult to implement optical buffers. Furthermore, the requirement for fast header processing, and strict synchronization makes OPS impractical using current technology.

OBS [3,4] was proposed as a new paradigm to achieve a practical balance between coarse-grained circuit switching and fine-grained packet switching. In OBS networks, incoming data is assembled into basic units, referred to as data bursts (DB), which are then transported over the optical core network. Control signaling is performed out-of-band by control packets (CP) which carry information such as the length, the destination address, and the QoS requirement of the optical burst. The control packet is separated from the burst by an offset time, which allows for the control packet to be processed at each intermediate node before the data burst arrives. OBS provides dynamic bandwidth allocation and statistical multiplexing of data, while having fewer technological restrictions than OPS. By aggregating packets into large sized bursts and providing out-of-band signaling, OBS eliminates the complex implementation issues of OPS. For example, no buffers are necessary at core nodes, headers can be processed at slower speeds, and synchronization requirements are relaxed in OBS. On the other hand, OBS incurs higher end-to-end delay and higher packet loss per contention compared to OPS, due to packet aggregation. Basic architectures for core and edge nodes in an OBS network have been studied in [5].

Packet Traffic (IP, ATM, etc.)			Periodic Traffic (SONET)
OPS	Connectionless OBS (JET, JIT, etc.)	Connection-Oriented OBS (TAW)	OCS (WR)
Optical Layer (DWDM)			

Figure 2. Supported services on optical networks.

Each of the three types of optical transport network architectures (OCS, OPS, OBS) may support different services, as shown in Fig. 2. Packet traffic can be supported by any of the three architectures in either a connectionless or connection-oriented manner. OPS and OBS support these different types of packet services through different signaling protocol implementations. In order to support connection-oriented services on OBS, a two-way reservation protocol such as, TAW can reserve the end-to-end path for the requested duration, prior to data transmission. Connectionless services on OBS can be supported by various one-way reservation protocols, such as, JET, JIT, and TAG [3, 4, 6]. Similarly, OPS may support connectionless services by routing packets on an individual basis, and may support connection-oriented services by assigning packets to flows and switching the flows based on labels applied to the packets. OBS differs from OPS primarily in that the signaling is done

out-of-band in OBS networks, while signaling is done in-band via packet headers in OPS networks. OCS supports packet traffic by establishing a logical topology consisting of lightpaths, and then switching or routing packets electronically over this logical topology. Signaling for establishing lightpaths in OCS networks is typically done out-of-band. In this article, we will focus on the connectionless mode of operation of OBS; however, the framework for the control plane will be general enough to support any out-of-band signaling scheme, including those for establishing OCS lightpaths.

An OBS network architecture can be represented in a layered manner as a set of protocols that provide services and exchange data with one other. A well-defined architecture with well-defined interfaces between the layers is essential for the practical implementation of OBS, as well as for the inter-operability of OBS with other networks. Furthermore, the layered hierarchy representation can provide a detailed insight into various implementation techniques, specifications, and functionalities of an OBS network. This article attempts to provide a layered view of different OBS protocols as well as some brief insight into the functionality of each protocol.

This article is organized as follows: Section 2 discusses the layered architecture of IP-over-OBS. Section 3 describes each layer of the OBS layered architecture, separating them into a data plane and a control plane. Section 4 provides a layered view of an OBS network, illustrating an end-to-end transmission to show what role each layer plays in the data transmission. Finally, Section 5 concludes the article.

2   IP-over-OBS Layered Architecture

An important objective in the design of OBS networks is the large-scale support of different legacy services, as well as emerging services. In this article, without loss of generality, we will discuss the OBS network as it supports IP traffic; however, the OBS architecture described here is general enough such that it is capable of supporting most types of higher-layer traffic.

Figure 3 shows the layered hierarchy of an IP-based OBS network. We call this hierarchy the IP-over-OBS architecture. In this representation the IP layer treats the OBS as its link layer, while the OBS operates on top of the optical (or DWDM) layer [7]. Thus, as a data transport system, the OBS network architecture implements the lower 3 layers, namely, physical, data link, and network layer.

Figure 4 shows our proposed OBS layered architecture, which follows the OSI reference model. In this representation we separate the control plane functionalities and protocols from those of the data plane. Such separation appears natural since the control information is transmitted out-of-band in OBS networks. Note that, in this model, we are ignoring the management plane, since the management plane communicates with all other layers and has no hierarchical relationship with them.
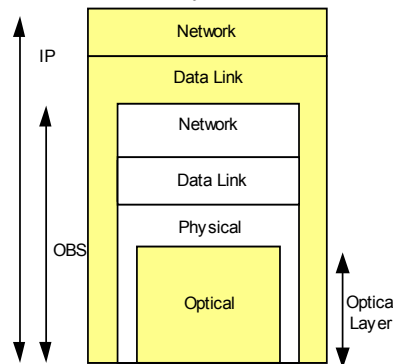

Figure 3. IP-over-OBS hierarchical layered architecture.

The control plane is responsible for transmitting control packets (CPs) while the data plane constructs and processes the data bursts (DBs). The CPs contain the information necessary for switching and routing DBs across the OBS network. The CPs are used for establishing the proper path prior to the arrival of the corresponding DBs, which arrive after some offset time. The CPs can also provide network management signaling.
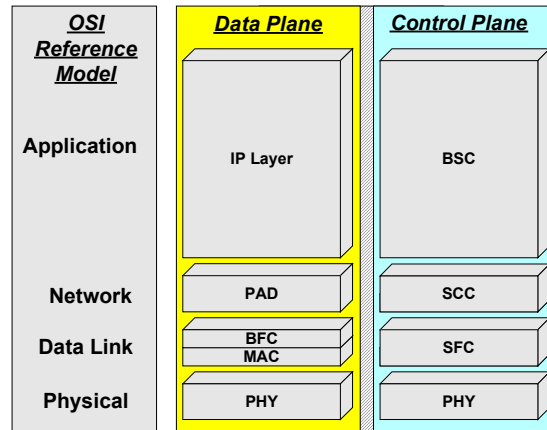

Figure 4. IP over OBS layered architecture.

Having two distinct planes suggests that each plane can operate independently of the other, using its own layers and protocols. Thus, it is conceivable to imagine that the DBs and CPs are encoded and routed on different transmission media.

In general, the OBS data plane architecture must take full advantage of DWDM technology and must support high capacity data transport links with no optical-to-electrical conversions. On the other hand, the design objective in the control plane is to make it flexible with low complexity. One way to achieve this goal is by processing CPs electronically. This approach offers high flexibility but limited processing capacity (a few tens of gigabits per second). Thus, simple encoding techniques and short frame lengths with minimum control overhead are required to allow fast and efficient CP processing. Transmitting CPs free of contention and in a highly reliable manner is also critical, since any error or loss of CPs results in higher data burst loss.

3 OBS Layered Architecture

In the following sections we describe basic functionalities of each layer in the data and control planes. We start with the data plane, which interconnects the OBS network with other client networks. For clarity, we describe the layered architecture of each plane in an order consistent with packet flow.

3.1 Data Plane Layers

The data plane transports incoming packets from the source edge node to a single or multiple destination nodes. Line cards in the edge node provide an interface with packets arriving from various client networks. The line cards can perform error detection and error correction on incoming IP packet headers. Since in this article we only consider IP-based OBS networks, we assume that all packets entering and leaving the OBS network are IP packets, and that these packets maintain their original format and structure.

3.1.1   Packet Aggregation and De-aggregation (PAD) Layer

The PAD layer aggregates incoming IP packets of same properties into data bursts. This layer also de-aggregates received data bursts into individual IP packets and assigns the packets to the proper outgoing links.

Transmitting IP packets at the ingress of an OBS network requires determining individual packet properties and aggregating the packets together. Packet properties include packet Quality-of-Service (QoS) and its client destination address. After each incoming IP packet is decoded, its destination address must be translated to an OBS equivalent edge node address. Packets with similar properties are then aggregated to form the burst payload.

The most common burst assembly techniques are timer-based and threshold-based. In timer-based burst assembly approaches [8], a burst is created and sent into the optical network at periodic time intervals; hence, the network may have variable length input bursts. In threshold-based burst assembly approaches [9], a limit is placed on the maximum number of packets contained in each burst. Hence, fixed-size bursts will be generated at the network edge. A threshold-based burst assembly approach will generate bursts at non-periodic time intervals. A combination of timer and threshold-based approaches has been proposed in order to reduce the variation in the burst characteristic due to the variations of load. In addition, a composite burst assembly approach [10] can be adopted in order to support QoS. A composite burst is created by combining packets of different classes into the same burst. The packets are placed from the head of the burst to the tail of the burst in order of decreasing class.

In the egress path, the PAD disassembles data bursts into IP packets. Each packet's header must be processed for its destination address and the type of service it requires. The destination address is translated to identify which line card the IP packet must be sent to. Line cards, in turn, forward packets to the appropriate interfaced client network such as a LAN or WAN.

The PAD layer contains various flow control mechanisms and offers sequence verification of incoming data bursts. The flow control protocols can pace the rate at which DBs are placed on a link. If data burst deflection routing is allowed throughout the OBS network, then DB re-sequencing at the destination node may be required to ensure ordered delivery of IP packets.

Various protocols may be considered to perform address translations. Intelligent protocols can dynamically keep track of network configuration changes and support broadcasting transmissions. In addition, numerous admission control schemes have been proposed to address IP packet aggregation techniques. Packet aggregation may be based on a single or multiple packet properties such as destination, class, or flow [5,8,9,10]. On the other hand, aggregation size (or degree of aggregation) remains as an important issue. For example, for low priority data bursts, a greater degree of aggregation results in greater burst loss probability. While dealing with these concerns, such protocols must also reduce packet end-to-end delay and assure QoS without impacting the bandwidth efficiency.

3.1.2   Burst Framing Control (BFC) Layer

The function of the burst framing control layer is to receive the aggregated packets from the higher layer (PAD) and to encapsulate them into proper frame structures. This layer also decodes incoming data burst frames and extracts the data field. Figure 5 represents a generic framing format of a data burst. Data burst frames have two characteristics: they can have variable length, and, by nature, they are non-periodic, meaning that they can arrive at any random time. Therefore, a framing pulse is necessary to indicate the beginning of each optical data burst frame. Framing pulses are typically isolated from the data-field by using a preamble to ensure data integrity.

Guard bands are normally a stream of fixed pulses used to separate consecutive frames. They are mandatory for reasons such as link length error, precision of clock distribution, and thermal effects. The checksum field may be required when data burst retransmission from the source to destination edge nodes is supported. In this case, edge nodes must be designed with considerable storage capacity. Use of the checksum may be considered especially when the medium does not offer the required transmission error rate.

In order to support burst segmentation [11], the data burst can consist of many individual frames, each frame containing one or more packets, represents a segment. The format of each segment is similar to that of a burst frame.
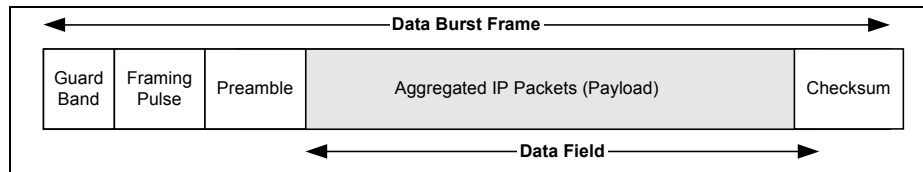

Figure 5. OBS framing structure of the data burst.

3.1.3   Medium Access Control (MAC) Sublayer

The MAC sublayer in data plane includes the reservation and scheduling protocols, the contention resolution schemes, and the offset time assignment protocols. The MAC layer can also provide class differentiation in order to provide higher protection for DBs with QoS requirements. The actual signaling process by which a node requests the network to setup or release a connection is performed in the control plane.

An OBS network is inherently a point-to-point network in which adjacent nodes are interconnected to each other through direct physical links. However, asynchronous data bursts entering a core node from different links may need to access the same outgoing link. The MAC sublayer provides a way to control access to the outgoing links among these data bursts. In general, access control schemes proposed for OBS networks can be categorized as *centralized* or *distributed*.

In a *centralized* OBS network [12] a single node (called the request server) will be in charge of data burst transmission throughout the entire network. Clearly, this mode of operation makes medium access straightforward since the request server provides a single point of coordination that eliminates contention and packet loss. However, centralized scheme is very complex and considered to have low reliability and robustness.

In a *distributed* OBS network each node operates autonomously. This scheme suffers from lack of any centralized coordination. Consequently, the number of DBs entering a node and attempting to access the medium may exceed the number of available channels of the outgoing port. This is the primary source of contention in distributed OBS networks. Therefore, efficient and reliable algorithms in the MAC sublayer are required to simultaneously minimize contention as well as expected end-to-end delay of DBs.

Based on the type of service requested by an application, such as connection-less or connection-oriented services, the OBS MAC needs to assign sufficient bandwidth and resources. Such assignments are obtained through the appropriate reservation protocols. Reservation protocols indicate the mechanisms in which a burst allocation starts and ends. Various out-of-band reservation approaches have been proposed for OBS networks. The most widely considered examples of such schemes are the Just-In-Time (JIT) reservation scheme (also know as tell-and-go), and the Just-

Enough-Time (JET) reservation scheme. Although the JET reservation scheme provides a more efficient use of bandwidth, its implementation requires higher complexity. Different variations of the JIT reservation scheme have been described in [6].

Various scheduling disciplines can be implemented in the MAC depending on the reservation protocols employed in the system. An OBS scheduling discipline determines the manner in which available outgoing data channels are found for DBs. Scheduling algorithms must be fast and efficient in order to lower the processing time and to minimize the data burst loss. Current data channel scheduling algorithms include first-fit unscheduled channel (FFUC) [5], latest available unscheduled channel (LAUC) or Horizon Scheduling [4,5], and latest available unscheduled channel with void filling (LAUC-VF) [5].

FFUC and LAUC scheduling disciplines can be considered for JIT whereas LAUC-VF is more practical for the JET reservation scheme. Note that scheduling protocols in the MAC layer should support class differentiation and provide a greater degree of protection and transmission reliability for high priority data bursts.

A major concern in distributed medium access control scheme is high contentions. Packet transmissions in this scheme can only be statistically guaranteed. Many different techniques and algorithms have been introduced to improve OBS reliability and to reduce the DB drop ratio. In general, such solutions can be divided into four basic categories: space deflection (such as deflection routing), time deflection (such as buffering and delaying the data) [13], wavelength conversion, and soft-contention resolution policies [11].

In addition to addressing scheduling algorithms and contention resolution policies, another function of the MAC sublayer is offset assignment and maintenance between DBs and their CPs. The offset time can be variable or fixed. However, as the CPs are processed and reconstructed at each hop, the offset times tends to be reduced. Such variations can result in miscalculating the DBs' arrival times. The MAC protocols dealing with these issues are referred as offset control protocols.

The MAC sublayer can also support establishing multipoint multicast connections. In these schemes any edge node can transmit its DBs to multiple destination edge nodes. Efficient signaling protocols can be implemented in the application layer of the control plane to support multipoint multicast.

3.1.4   Physical (PHY) Layer

The physical layer of OBS is responsible for the actual transport of DBs and CPs from one node to another. It includes converting signals into appropriate electrical or optical format and uploading DBs into appropriate transmission frames. The physical layer also defines the actual physical interfaces between nodes in OBS. The PHY is divided into two sublayers [15]:
-    Data transport component,
-    Medium dependent component.
We describe these sublayers briefly in the following paragraphs.

Data Transport Component (DTC): This is the medium independent sublayer of the physical layer. In the ingress direction it encodes data bits into specific pulse transmission called line codes (such as NRZ, AMI, HDB3, etc) and performs electrical/optical conversions. This sublayer also specifies transmission capacity.

Furthermore, DTC is responsible for implementing mechanisms to resolve synchronization issues between nodes including transmission techniques (slotted or unslotted). For example, in an unslotted asynchronous network where each node has its own internal clock, DTC ensures sufficient inter-

frame gap and defines the maximum allowable clock variation. The DTC also specifies the buffering requirement to alleviate any clock jitters among nodes.

Medium Dependent Component (MDC): This sublayer deals with the actual type of the medium used to transmit CPs and DBs including, coax, radio frequency, or optical fiber[1]. Selections of connectors, transmitters, receivers, etc., are considered as parts of the MDC sublayer. In an OBS network, as a special category of burst switching, the MDC is transparent to the photonic (WDM) sublayer, which provides lightpaths to the network. A lightpath is an end-to-end connection established across the optical network, and the lightpath uses a wavelength on each link in the path between the source and destination. Consequently, tasks such as optical amplification and wavelength conversion are defined in the MDC sublayer.

3.2   Control Plane Layers

We now turn our attention from the data plane to the control plane. As we mentioned earlier, separation of planes in the OBS network architecture was inspired by the need to provide practical and reliable medium access protocols at high speeds. Due to current technological limitations in all-optical packet switching, it is not practical to implement MAC protocols in the data plane without interrupting data by optical-electrical converters. In OBS networks, implementing the MAC sublayer as the application layer of the control plane allows arbitration protocols to be performed in a domain (electrical) independent of data (optical).

The control plane in an IP-centric OBS network can be based on existing protocol standards. For example, similar to the Internet protocol, we can implement the Resource reSerVation Protocol (RSVP) and the Open Shortest Path First (OSPF) protocol in the control plane to provide signaling. Such standard protocols support a variety of control functionalities as well as multipoint multicasting. However, the major issue with implementing such protocol structures is their complexity and long processing time requirements. With this motivation, new protocols maybe considered to optimize control message processing and new signaling semantics. Key features of the new protocols must be flexibility and low complexity. In the following paragraphs we briefly describe general signaling semantics and basic functionalities of each layer in the control plane.

3.2.1   Burst Signaling Control (BSC) Layer

The BSC layer contains the data plane MACs' scheduling, contention resolution, and offset control protocols through its signaling protocols. Data burst properties including destination address, quality-of-service, etc., are passed to the BSC layer from the MAC sublayer. The BSC layer determines the type of the control packet to be transmitted to the next hop. Typical examples of the control packet types are burst header packets (BHP), burst cancellation packets (BCP), or network management packets (NMP). BHPs contain their associated data burst properties, BCPs can be used to cancel an existing reservation in downstream nodes, and NMPs provide network status information. Other types of control packet can be considered to support multipoint multicasting connections.

Each received control packet on the incoming port is identified by its type and its BSC functions accordingly. For example, if a BHP is received, its data burst reservation request is checked for adequate bandwidth and, upon verification of availability, the request is scheduled. New changes in data burst reservations must be communicated to the switch fabric control unit to update its scheduling table.

---

[1] Note that the concept of burst switching can be implemented on various mediums and it is not limited to optical burst switching.

3.2.2 Signaling Connection Control (SCC) Layer

Similar to the PAD layer in the data plane, the SCC layer includes the routing algorithms for control packets in order to establish the physical path for incoming data bursts. Hence, the actual data burst routing also takes place in this layer. Note that, in general, since the data and control planes can be implemented on separate mediums, it is possible that the physical routing paths for CPs and DBs are different. Various routing protocols can be considered for implementation in the SCC layer. A rough taxonomy to categorize routing protocols is presented in [16].

3.2.3 Signaling Frame Control (SFC) Layer

The main purpose of the SFC layer is to provide reliable transmission of control packets. The SFC layer can be considered as a pure data link protocol operating between adjacent nodes. The SFC layer receives bit streams containing the control packet type and its associated data burst properties, and it constructs CP frames by attaching overhead bits. Many popular framing mechanisms such as High-Level Data Link Control (HDLC) may be considered for the data link protocol. However, the protocol complexity and cost are critical as interface speed increases. Figure 6 shows a generic framing format of a control packet frame.
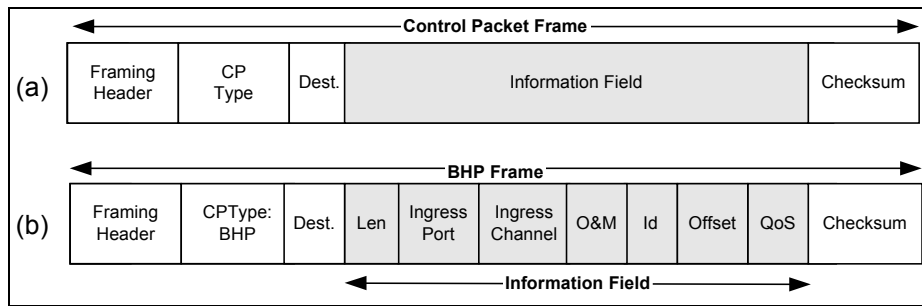

Figure 6. OBS framing structure of the control packet; (a) a generic control packet frame, (b) BHP frame

Typically, control packets in OBS networks are continuous fixed-size packets, which are processed electronically. Therefore, there is no need for attaching a framing pulse and for the use of a preamble. However, each CP must still contain its own framing header.

To guarantee fast processing of control packets at each node, CPs must contain limited information, yet, it is crucial to protect control packets from errors on each link. Transmission errors in control packets can result in bits being changed in the information field. Incorrect bits will be misinterpreted in the downstream core node and result in, for example, dropping high priority bursts, incorrect switch fabric setup, or even burst misrouting. To protect the CP from error, a cyclic redundancy check (CRC) can be implemented in the checksum field. CRC codes can provide a large selection of error correcting capacity [15]. Each CP must also have a destination field indicating its destination node. Furthermore, all CPs must have a type indication specified in the CP-type field. Different CP types were described in the Burst Signaling Control section. Contents of the information field vary depending on the CP type.

If a control packet is associated with an incoming DB, it is referred to as a BHP. A typical BHP frame is shown in Figure 6 (b). Note that the BHP information field is divided into several fields including length, ingress port, and ingress channel, which refer to DB's duration, its edge node source, and the wavelength on which it is expected to arrive, respectively. The id field can be useful for checking data burst sequencing when deflection routing is allowed. The QoS and offset fields indicate the incoming data burst priority level and the offset time between a BHP-type control packet and its associated data burst, respectively. The O&M field contains network management related signaling information, such as loop-back requests, protection switching, or link failure notification.

3.2.4 Physical (PHY) Layer

The physical layer in control plane performs similar functionalities to the data plane's physical layer but it may have different characteristics. One such difference is the transmission rate. Typically, control packets are transmitted at lower rate than data bursts in order to achieve practical packet processing. CPs' transmission rate and offset time are generally designed for optimum performance in terms of end-to-end delay and bandwidth efficiency.

In addition, as in the data plane, the PHY layer addresses synchronization issues and determines transmission techniques such as slotted or unslotted transmission. Since control packets are processed electronically, they can be used to retrieve an adjacent node's timing information. Such hop-by-hop clock recovery schemes allow the network to synchronize to a single reference clock.

4 Layered View of an OBS Network

In this section we attempt to summarize our proposed OBS layered hierarchy by means of an example. Figure 7 shows various transport stages in a simple OBS network configuration where DBs and their control packets are transmitted on the same link but on different wavelengths and are separated in time by an offset. We assume that each link contains a single control channel, the core node is bufferless, and data bursts are IP-based.

Stage 1 – IP Interface: The edge node's line cards receive IP packets from various client networks. They process packet headers and extract the type of service and the destination client address. The aggregation protocol in the PAD layer translates the destination client address into an OBS edge node address and determines the next hop. The protocol also classifies packets based on their destination, type of service, or both.

Stage 2 – Burst Aggregation: The PAD layer aggregates IP packets of the same class into bursts. An admission control protocol examines bursts and verifies whether or not they are ready for transmission. Upon completion of the aggregation process, bursts will be passed on to the BFC layer. The BFC encapsulates the aggregated IP packets into data burst frames and places them into the DB buffers. DBs are stored until the MAC sublayer assigns the DBs to a transmission time slot. The MAC sublayer sends the data burst properties to the BSC layer in the control plane. The BSC, in turn, constructs a control packet of type BHP and passes it to the SCC layer. The SCC layer assigns a routing path based on the DB destination address.

The BHP is handed to the SFC layer to be placed in the proper framing format. The BHP frame is stored in the BHP buffer and waits for the SFC to search for the next available time slot on the control channel. Once a potential BHP time slot is found, the SFC is required to verify the bandwidth availability on data channels based on the offset value (this is only required if the offset is not fixed). Assuming that the scheduling was successful, the BHP frame is passed on to the physical layer and transmitted on the link. After an offset time, the DB is sent to the physical layer on the data plane and transmitted on the pre-assigned channel.

Stage 3 – Burst Switching: The core node receives the BHP and processes it electronically. The SFC verifies the BHP's checksum and, assuming the checksum matches the calculated value, extracts information fields from the frame. The SCC identifies the requested connection in the core node and its duration. This reservation request is sent to the BSC layer and verifies available bandwidth and decides to either make the reservation or to discard it. If the reservation was successful then the reservation table in the switch fabric control unit will be updated. The expected DB arrives after the offset time and it cuts-through the pre-established path in the optical switch. Thus, the DB only goes through the physical layer of the data plane without any O/E interruptions.

Similar steps as described in the *burst aggregation* stage take place in order to transmit the control packet on the outgoing control channel. The BSC creates a new control packet of type BHP and the SCC assigns the next hop's outgoing port. The CP is encapsulated into a frame and transmitted on the assigned time slot. On the other hand, data bursts only go through amplification and wavelength conversion, which may be required for compensating power loss and accessing the outgoing port, respectively. The loss of power occurs as optical signals are decoupled and traverse through the optical switch fabric

Note that, as shown in Figure 7, by processing the CPs in the control plane, we can essentially implement the data plane MAC sublayer without having to process DBs.

Stage 4 – Burst Disassembly and IP Forwarding: At the destination node all CPs and DBs will be terminated and processed electronically. The CPs are verified for errors in the SFC layer and upon detecting any errors, their associated DBs will be discarded (consequently, the destination edge node may request retransmission). Similarly, data channels are de-multiplexed and DBs are verified for transmission errors and de-framed in the BFC layer. The data burst payload is passed to the PAD layer and decomposed into individual IP packets. This layer can also check for DB order and decide what action to take (such as buffering or discarding) with out-of-sequence DBs.

Disassembled IP packets are processed to identify the client network to which they should be forwarded. The PAD layer needs to translate these addresses and to determine the destination line card. The line card, in turn, forwards the packets to the proper client networks.
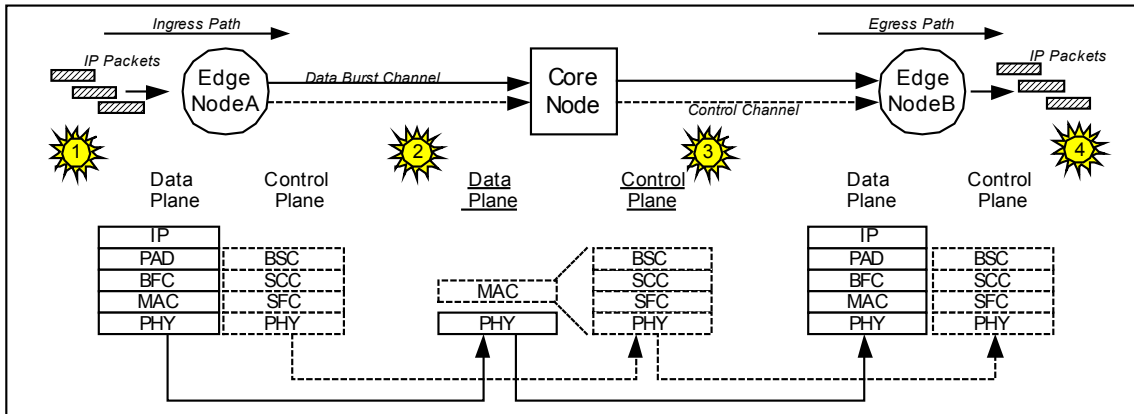

Figure 7. Layered view of an OBS network.

Table 1 summaries the basic functionalities of data as well as control plane in the proposed OBS layered architecture. The references addressing these functionalities are mentioned in the table.

Table 1. Summary of OBS layer functionalities.

Plane	Description	Reference
Data Plane	**IP Layer:** • Transmits and receives IP packets	[16]
	Packet Aggregation and De-aggregation (PAD): • Classifies packets based on their properties such as, QoS requirements and destination address • Aggregates IP packets into burst payload • Translates IP address to OBS MAC address • Disassembles bursts into individual packets • Verifies data burst sequence id • Controls data burst transmission flow	[8], [9], [10]
	Burst Framing Control (BFC): • Encapsulates the burst payload into a burst frame	
	Medium Access Control (MAC): • Provides methods to access outgoing ports • Assigns sufficient bandwidth and resources for each request • Handles scheduling and contention resolution schemes through the control plane. • Supports multicasting	[3], [4], [5], [6], [11], [12]
	Physical Layer (PHY): • Determines transmission rate and capacity • Specifies transmission technique (slotted, unslotted) • Converts signals between optical and electronic domains • Handles synchronization schemes and clock recovery issues	[8], [14], [15]
Control Plane	**Burst Signaling Control (BSC):** • Implements scheduling and contention resolution protocols • Generates control packets • Implements the signaling scheme	[3], [4], [5], [6], [11], [12]
	Signaling Connection Control (SCC): • Assigns routing and forwarding protocols	
	Signaling Frame Control (SFC): • Implements the hop-by-hop data link protocols • Constructs the control packet frame	
	Physical Layer (PHY): • Determines transmission rate and capacity • Specifies transmission technique (slotted, unslotted) • Converts signals between optical and electronic domains • Handles synchronization schemes and clock recovery issues	[8], [14], [15]

5   Summary

Optical burst switching has been proposed as a practical approach for supporting the next-generation high-speed high-capacity Internet. A layered architectural representation of the OBS network can be used as a baseline for understanding protocol requirements as well as their future development and design. In this article we provided an organized decomposition of the different layers for supporting OBS networks. Detailed descriptions of each layer along with their functionalities and related protocols were presented. To furnish a better understanding of the proposed layered architecture, an illustrative example of an end-to-end data transmission was provided. The proposed layered

architecture can be used as a baseline for future development and design of protocols and interfacing functions over optical burst-switched networks.

6 References

[1] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath Communications: An Approach to High Bandwidth Optical WANs," IEEE Transactions on Communications, vol. 40, no. 7, pp.1171-1182, 1992.

[2] D. K. Hunter et al., "WASPNET: A Wavelength Switched Packet Network," *IEEE Communication. Magazine,* Mar. 1999, pp. 120-129.

[3] C. Qiao and M. Yoo, "Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet," *Journal of High Speed Networks*, vol. 8, no.1, pp. 69-84, Jan. 1999.

[4] J.S. Turner, "Terabit burst switching," Journal of High Speed Networks, Vol. 8, No. 1, January 1999, pp. 3-16.

[5] Y. Xiong, M. Vanderhoute, and H.C. Cankaya, "Control Architecture in Optical Burst-Switched WDM Networks," IEEE Journal on Selected Areas in Communications, Vol. 18, No. 10, October 2000, pp. 1838-1851.

[6] L. Xu, H.G. Perros, and G. Rouskas, "Techniques for Optical Packet Switching and Optical Burst Switching," IEEE Communications Magazine, vol. 39, no. 1, January 2001, pp. 136-142.

[7] R. Ramaswami and K.N. Sivarajan, "Optical Networks: A Practical Perspective," Morgan Kaufmann Publishers, 1998.

[8] A. Ge, F. Callegati, and L.S. Tamil, "On Optical Burst Switching and Self-Similar Traffic," *IEEE Communications Letters*, vol. 4, no. 3, pp. 98-100, March 2000.

[9] V.M. Vokkarane, K. Haridoss, and J.P. Jue, "Threshold-Based Burst Assembly Policies for QoS Support in Optical Burst-Switched Networks," *Proceedings, SPIE OptiComm 2002*, Boston, MA, vol.4874, pp. 125-136, July 2002.

[10] V.M. Vokkarane, Q. Zhang, J.P. Jue, and B. Chen, "Generalized Burst Assembly and Scheduling Techniques for QoS Support in Optical Burst-Switched Networks," *Proceedings, IEEE Globecom 2002,* Taipei, Taiwan, November 2002.

[11] V. Vokkarane, J.P. Jue, and S. Sitaraman, "Burst Segmentation: an Approach for Reducing Packet Loss in Optical Burst Switched Networks," *Proceedings, IEEE ICC 2002*, New York, NY, vol. 5, pp. 2673-2677, April-May 2002.

[12] E. Kozlovski, M. Duser, I. de Miguel, and P. Bayvel, "Analysis of Burst Scheduling for Dynamic Wavelength Assignment in Optical Burst-Switched Networks," IEEE LEOS 2001, vol. 1, pp. 161-162, 2001.

[13] S. Yao, B. Mukherjee, S.J.B. Yoo, and S. Dixit, "All-Optical Packet-Switched Networks: A Study of Contention Resolution Schemes in an Irregular Mesh Network with Variable-Sized Packets," *Proceedings, SPIE OptiComm 2000*, Dallas, TX, pp.235-246, Oct. 2000.

[14] R. Jain, "FDDI Handbook: High-Speed Networking with Fiber and Other Media," Addison-Wesley, Reading, MA, April 1994.

[15] S. Keshav, "An Engineering Approach to Computer Networking: ATM Networks, the Internet, and the Telephone Network," Addison-Wesley, 1997.

[16] A.S. Tanenbaum, "Computer Networks," Third Edition, Prentice Hall, 1996.