

Transporting IP Packets over Light: A Survey

Lisong Xu, Harry Perros, George Rouskas

Department of Computer Science
North Carolina State University
Raleigh, NC 27695-7534
{lxu2, hp, rouskas}@eos.ncsu.edu

Abstract

Optical networks appear to be the solution of choice for providing a faster networking infrastructure that can meet the explosive growth of the Internet. One of the key issues that needs to be addressed is how optical networks can be used to carry IP traffic. In this paper, we survey some of the techniques that have been proposed in the literature for transporting IP traffic over an optical network. Specifically, we survey techniques for IP over SONET, optical packet switch architectures, and optical burst switching.

1. Introduction

The explosive growth of the Internet has given rise to the need for faster transfer techniques. Optical transport networks appears to be the solution of choice. One of the key issues that needs to be addressed is how can optical networks be used to carry IP traffic. In this paper, we review some of the techniques that have been proposed in the literature for *IP over light*, that is, for transporting IP traffic over an optical transport network.

SONET/SDH and FDDI networks are considered as *first-generation optical networks*. They are single-wavelength optical networks and switching is done electronically. That is, at each switch, each incoming optical signal is converted to the electrical domain, the traffic is switched electronically and then the signal is converted back to optical and transmitted out. IP traffic can be transmitted over SONET using the following two combinations of encapsulation: IP/ATM/SONET and IP/PPP/HDLC/SONET. Recently Lucent proposed a new model IP/SDL/SONET [10].

Wavelength-division multiplexing (WDM) is the enabling technology for the *second-generation optical networks*. In WDM, multiple wavelengths can be used in a single fiber to transmit information. WDM-based networks are referred by the ITU-T as *optical transport networks*. There are several different technologies that have been developed in order to enable the transfer of data over WDM, such as wavelength routing, broadcast-and-select, optical packet switching, and optical burst switching. Wavelength routing networks have already been deployed. Also, broadcast-and-select networks have been extensively studied and several prototypes have been developed. Optical packet switching and optical burst switching are still in the research phase. These technologies can support IP traffic, as well as ATM traffic.

Wavelength routing is a form of optical circuit switching, where a dedicated connection is established. This connection is known as *lightpath* and it consists of the same wavelength allocated on each link along the path. The lightpath may consist of different wavelengths along the path if converters are present. Usually there is almost no requirement for the format and the rate of the data transmitted in the lightpath. Data remains as an optical signal throughout the entire lightpath. In view of this, it is often referred to as a *transparent* lightpath. A first-generation optical network can run on top of a transparent lightpath. This technology is not described in this paper. The interested reader is referred to [1] and [2] for further details.

An optical packet network consists of optical packet switches interconnected with fiber running WDM. The switches may be adjacent, or they may be connected by lightpaths. Several different designs of optical packet switches have been recently proposed in the literature. The user data is transmitted in optical packets, which are switched within each optical packet switch entirely in the optical domain. Thus, the user data remains as an optical signal in the entire path from source to destination. No optical-to-electrical or electrical-to-optical conversions are required. Optical packet switches are discussed in detail in Section 4.

Optical burst switching is a technique for transmitting bursts of traffic through an optical transport network by setting up a connection and reserving resources end-to-end for only one burst. This technique is an adaptation of an ITU-T standard for burst

switching for ATM networks, known as *ATM block transfer* (ABT). There are two versions of ABT, namely, *ABT with delayed transmission* and *ABT with immediate transmission*. In the first case, when a source wants to transmit a burst, it sends a packet to the ATM switches along the path of the connection to inform them that it wants to transmit a burst. If all the switches can accommodate the burst, the request is accepted and the source is allowed to go ahead with its transmission. Otherwise, the request is refused and the source has to send another request later on. In ABT with immediate transfer, the source sends the request packet, and then immediately after it, without receiving a confirmation, transmits its burst. If a switch along the path cannot carry the burst, due to congestion, the burst is dropped. These two techniques have been adopted for optical networks.

The paper is organized as follows. In following section, we give a brief overview of SONET and describe the various techniques for transporting IP packets over SONET. Section 3 gives an overview of how IP packets can be transported over WDM. In Section 4, we give various optical packet switch architectures that have been proposed in the literature. In Section 5, we describe optical burst switching, and finally Section 6 contains the conclusions.

2. IP over SONET

IP over SONET is typically of interest to large-scale *Internet Service Providers* (ISPs), who operate SONET links in their backbone networks. In many cases, ATM runs on top of SONET, and then IP traffic is transported by the ATM network using classical IP and ARP over ATM. This transfer mode of IP traffic is referred to as IP/ATM/SONET. However, the ATM layer introduces an additional overhead, due to the ATM header. In view of this, an alternative scheme has been introduced that relies on PPP and HDLC, referred to as IP/PPP/HDLC/SONET. Recently, IP/SDL/SONET has been proposed to support IP traffic. SDL uses a different framing mechanism that has a good scalability in high-speed networks and has predictable transmission overhead. In this section, we first review briefly SONET, and then we discuss the above three schemes in detail.

SONET stands for *synchronous optical network*, and it was developed by Bellcore (now Telcordia) for the optical fiber backbone of the telephone network. An outgrowth of SONET was standardized by ITU-T, and it is known as the *synchronous digital hierarchy* (SDH).

The basic SONET frame is the Synchronous Transport Signal-1 (STS-1), which can be visualized as a two-dimensional matrix of 9 rows by 90 columns bytes. The first 3 columns are transport overhead and they contain the section overhead and the line overhead. The remaining columns, i.e. the fourth to the last column, are called the *synchronous payload envelope* (SPE). However, not all the SPE is used by the end user. Column 4 is the path overhead. Column 30 and 59 are called the fixed-stuff and they do not contain any information. The optical signal transferred by STS-1 is referred to as Optical Carrier-level 1 (OC-1). N STS-1 frames can be multiplexed by a time division multiplexer to produce an STS- N frame. An STS- N frame has 9 rows and $90N$ columns. The first $3N$ columns are transport overhead. The path overhead contains N columns from column $3N+1$ to column $4N$.

If N STS-1 frames have the same source and the same destination, they will be multiplexed to produce one STS- Nc frame where c stands for concatenated frames. The significant difference between an STS- N frame and an STS- Nc frame is in the number of path overhead columns and the number of fixed-stuff columns. An STS- N frame has N columns of path overhead, while an STS- Nc frame has only 1 column of path overhead. An STS- N frame has $2N$ columns of fixed-stuff, while an STS- Nc frame has $N/3-1$ columns of fixed-stuff. STS- Nc frames are often used to carry ATM cells and IP packets at higher rates. One common used format STS-3c is shown in Figure 1. It consists of 1 path column and $3/3-1$ fixed-stuff column. Thus, the payload available for carrying ATM cells or IP packets is $(9 \text{ rows}) \times (270-9-1 \text{ columns}) = 2340$ bytes. Note that regardless of the value of N , each STS- N frame or STS- Nc frame is transmitted within 125 microseconds. The rates of SONET links range from several Mbps to several Gbps.

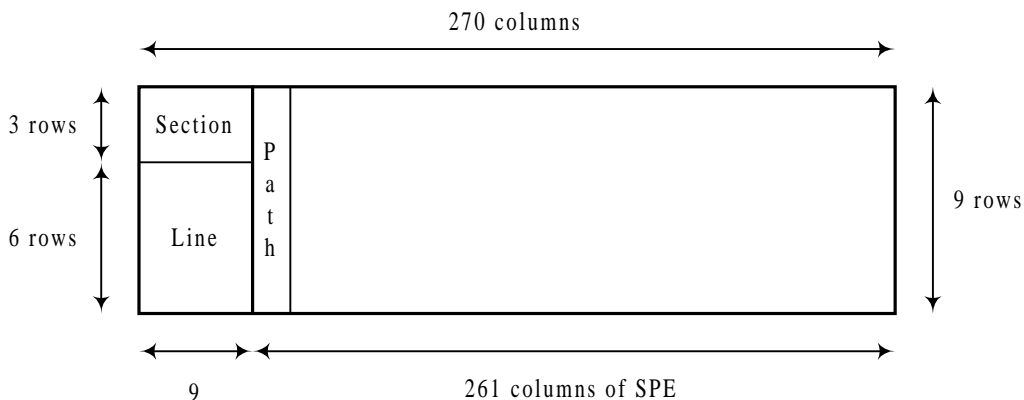


Figure 1: The STS-3c frame structure

A one-byte field in the path overhead of a SONET frame, referred to as C2, indicates the specific payload mapping. For example, if C2 is populated with 0x13, then the payload is ATM cells. If it is populated with 0xCF, then the payload is PPP frames in HDLC-like framing with no payload scrambling as specified in RFC 1619. 0x16 indicates that the payload is also PPP frames but with an ATM-style $x^{43}+1$ scrambler as specified in RFC 2615. The code 0x17 has been proposed to indicate that the payload is SDL frames.

Scrambling is performed to assure that there are enough transitions between 0's and 1's on the optical fiber. For, too many successive 0's or 1's will cause the loss of bit clocking information. With the exception of some overhead bytes, all bytes are scrambled. Bits are XOR'ed with a pseudo-random sequence which is generated by the polynomial $1+x^6+x^7$ with a seed of 1111111. One problem associated with this scrambler is that for some specified values of the user data, the result is all 0's or 1's.

SONET is typically deployed in a dual-ring topology. The dual-ring provides fault tolerance by switching from the working ring to the protecting ring when a fault occurs. Several kinds of equipment is used in SONET, such as, section terminating equipment (STE), line terminating equipment (LTE), and path terminating equipment (PTE). The primary function of the STE is to provide framing, scrambling, and error control. LTE is concerned with functions, such as protection switching, multiplexing and synchronization. PTE is concerned with the transport of various types of payload.

2.1 IP/ATM/SONET

In this scheme, IP packets are transferred over an ATM network, which consists of ATM switches interconnected by SONET links. This scheme is preferred by carriers which usually provide different services that require different types of quality of service. ATM networks provide a single platform for the transmission of diverse types of traffic, such as voice, video, and data. ATM can provide different quality of service to different applications at speeds varying from fractional T1 to gigabits/s.

2.1.1 Encapsulation of IP packets into ATM cells

An IP packet is encapsulated into an AAL 5 PDU, which is then fragmented into a sequence of 48-bytes segments, and each segment is carried in the payload of an ATM cell. Multiprotocol encapsulation over AAL 5 (RFC 1483) is used to encapsulate IP packets into AAL PDUs. There are two mechanisms, namely, *VC based multiplexing*, and *LLC Encapsulation*. The difference between them is that the former allows only one protocol to be carried in a single VC while the latter allows multiple protocols in one VC.

In VC Based Multiplexing, since only one protocol is carried by a single VC, IP packets are directly encapsulated into the payload field of AAL5, as shown in Figure 2. In LLC Encapsulation, an encapsulation header is added to each packet to indicate the information of the packet. For IP packets, an 8-byte header including an IEEE 802.2 LLC header and an IEEE 802.1a SNAP header is added before it is encapsulated into the AAL5 payload field, as shown in Figure 3. In both encapsulations, a PAD field with the length of no more than 48 bytes and an 8-byte trailer are created to build an AAL5 frame.

Information (IP Packets) 0-64KB	Padding 0-47 Bytes	AAL5 Trailer 8 Bytes
------------------------------------	-----------------------	-------------------------

Figure 2: AAL5 Frame with VC Based Multiplexing

LLC&SNAP Header 8 Bytes	Information (IP Packets) 0-64KB	Padding 0-47 Bytes	AAL5 Trailer 8 Bytes
----------------------------	------------------------------------	-----------------------	-------------------------

Figure 3: AAL5 Frame with LLC Encapsulation

2.1.2 Encapsulation of ATM cells into SONET frames

The ATM Forum has specified the mapping of ATM cells into various SONET frames. For STS-3c, ATM cells are directly mapped into the STS-3c frame row by row by aligning the byte of ATM cells with the byte of the STS-3c frame. As noted before, the available payload of STS-3c frames for carrying ATM cells is 2340 bytes which is not an integer multiple of the ATM cell (53 bytes). In view of this, it is possible that an ATM cell may lie over two consecutive SONET frames.

Because SONET expects ATM cells to arrive at a rate equal to the capacity of the SONET link, idle ATM cells or unassigned ATM cells (dummy cells) are inserted into SONET payload during the time that no ATM cells are available for transmission.

Moreover, before ATM cells are placed into the SONET payload, they are scrambled to prevent malicious attacks.

Delineation of ATM cells carried in the SONET payload is achieved by the header error control (HEC) byte in the ATM cell header. Using CRC the HEC field is used to detect single-bit or multiple-bit errors in the first 32 bits of the ATM header. The state diagram of delineation consists of three states: the hunt state, the presync state, and the sync state. The diagram starts at the hunt state. At this state, the delineation is processed byte by byte to calculate the HEC byte for the assumed header field. When a correct HEC is found, a header is assumed to be found and the state machine enters the presync state. At the presync state, HECs are sought on a cell-by-cell basis. If a wrong HEC is found, the state machine goes back to the hunt state. If six consecutive correct HECs are found, the state machine enters the sync state. At the sync state, delineation goes back to the hunt state if seven consecutive wrong HECs are found.

2.1.3 Transporting IP packets over ATM

Various schemes have been proposed to carry connectionless traffic, such as IP traffic, over ATM, which is connection-oriented. The IETF and the ATM Forum have proposed a number of solutions, such as *LAN emulation* (LANE), *classical IP and ARP over ATM*, and *next hop routing protocol* (NHRP). The development of these techniques was motivated by the desire to introduce ATM technology with as little disruption as possible to the existing IP network. In view of this, the ATM network is simply used as a fast datalink. LANE, as the name implies, emulates the characteristics and behavior of legacy LANs over an ATM network. It allows existing applications to run over an ATM network without any modifications. Classical IP and ARP over ATM was developed for a single IP subnet, that is, for a set of nodes that have the same IP network number and subnet mask. The members of the subnet communicate with each other directly over ATM connections. IP subnets are interconnected by conventional IP routers. Address resolution within the subnet is an important function of the protocol. This is necessitated by the fact that a host has both an IP address and an ATM address. IP addresses are different to ATM addresses. Thus, there is a need to translate the IP address of a host to its corresponding ATM address and vice versa. The next hop routing protocol (NHRP) is an address resolution technique for resolving IP addresses with ATM addresses in a multiple subnet environment.

An alternative technique for switching IP packets was introduced by Ipsilon Networks (which was later on purchased by Nokia), known as *IP switching*. Similar schemes were proposed by other companies, such as CISCO's *tag switching*. These techniques are known as *label swapping* techniques, because they utilize the label swapping mechanism of the ATM switch. They have been standardized into the *multiprotocol label switching* (MPLS).

In IP Switching, an IP switch consists of an ATM switch and an IP switch controller. An IP packet is segmented into ATM cells, which are forwarded hop-by-hop to the next IP switch controller over a default virtual connection. When the cells are received by the IP switch controller, they are reassembled into the original IP packet, and a IP routing decision is made. Subsequently, the IP packet is again segmented to ATM cells which are forwarded to the next IP switch. Great savings can be achieved if the routing decision is done only for the first IP packet, and then cashed into the ATM switch for

the remaining IP packets belonging to the same flow. (A flow of IP packets is a sequence of packets belonging to the same source/destination pair.) IP switching is a mechanism that capitalizes on this idea. Specifically, it permits subsequent ATM cells belonging to IP packets of the same flow to be switched through the IP switch without having to a) assemble them into the original IP packet, and b) make an IP routing decision.

2.2 IP/PPP/HDLC/SONET

In the above described solution, which is based on ATM networks, each IP packet is divided into a number of 48-byte segments, after it is encapsulated in an AAL 5 PDU. Each segment is then carried by an ATM cell. We recall that the ATM cell has a 5-byte header. In view of this, this scheme introduces an additional overhead due to the ATM header. An alternative solution that avoids this overhead, is to encapsulate IP packets into PPP frames with an HDLC-like framing and then transmit over SONET. This scheme was proposed in RFC 1662.

2.2.1 Encapsulating IP packets into PPP/HDLC frames

PPP, defined in RFC 1661, is a point-to-point data link layer protocol that has three functions: encapsulating multi-protocol datagrams; establishing, configuring, and testing the data-link connection; and, establishing and configuring different network-layer protocols.

Flag 0x7E	Address 0xFF	Control 0x03	Protocol 1/2 bytes	Information 0-64 KB	FCS 2/4 bytes	Flag 0x7E	Inter-frame Fill or Next Address
--------------	-----------------	-----------------	-----------------------	------------------------	------------------	--------------	-------------------------------------

Figure 4: The PPP frame with HDLC-like framing

The format of a PPP frame with HDLC-like framing is shown in Figure 4. Each frame begins and ends with a flag set equal to 0x7E. The ending flag of a frame can be used as the beginning flag of the next frame. Byte-stuffing, explained below, is used to guarantee that a byte in the payload equal to 0x7E is not interpreted as a flag. The address field is always set to the hexadecimal value 0xFF, which means that all stations are to accept this frame. The default value of the control field is 0x03, which indicates that frame sequence numbers are not used. The protocol field indicates what kind of packets are in the payload field. CRC is used to provide error control and the frame check sequence is saved in the FCS field, and it is calculated for all the other fields. The length of the protocol field and the FCS field can be variable, and it is determined by PPP at set-up time. RFC 1619 suggests a 16-bit protocol field and a 32-bit FCS field.

An IP packet is encapsulated in the information field of a PPP frame where the protocol field is set to 0x0021 for IPv4 and to 0x0057 for IPv6.

2.2.2 Encapsulation of PPP/HDLC frames into a SONET frame

PPP over SONET was defined in RFC 1619, and was subsequently updated in RFC 2615. PPP/HDLC frames are located row by row and byte-aligned within the payload of a SONET frame. A PPP/HDLC frame may lie between two SONET frames. During the time that there are no IP packet available for transmission, the SONET frame is

filled with flags (0x7E). At the receiver side, these flags are discarded. In RFC 1619, no scrambling is recommended for PPP over SONET, while an ATM-like scrambling is recommended in RFC 2615.

In RFC 2615, the order in which operations are carried out is described. Specifically, at the transmitter side we have the following order: encapsulation of IP packet in a PPP/HDLC frame, FCS calculation, byte stuffing, scrambling, and placing of the resulting frame in the SONET payload. At the receiver side, we have the opposite order of operations: extraction of the frame from the SONET payload, descrambling, byte destuffing, CRC check, and finally extraction of the IP packet from the PPP/HDLC frame.

Byte-stuffing is used to make sure that a byte in the payload equal to 0x7E is not interpreted as a flag. At the transmitter side, the payload field of PPP frames is monitored for the flag byte (0x7E) and the escape byte (0x7D). The flag byte is replaced by two bytes 0x7D 0x5E. The escape byte is replaced by two bytes 0x7D 0x5D. At the receiver side, the original bytes are recovered.

Delineation of frames within the SONET payload is based on the flag field. The receiver continuously checks for the flag. The bytes between two flag bytes make up a PPP frame. However, two consecutive flags constitute an empty frame, which is discarded.

2.3 IP/SDL/SONET

The Simple Data Link Protocol (SDL) proposed in [10] uses the length-based method to delineate frames, instead of using the flag-based method as in HDLC. Byte-stuffing used in flag-based delineation creates variable-length transmission overhead for different payloads. For example, in IP/PPP/HDLC/SONET, if an IP packet contains x number of bytes with the value of 0x7E, the HDLC transmission overhead is at least x bytes. It becomes more difficult to accurately calculate the packet delay. Length-based delineation creates constant transmission overhead for different payloads. For example, in the basic mode of SDL, no matter whatever an IP packet contains, the SDL transmission overhead is always 4 bytes per SDL frame. From this point of view, SDL is more suitable for the next generation of QOS-capable networks than HDLC.

2.3.1 Encapsulation of an IP packet into an SDL frame

The SDL frame format is shown in Figure 5. The length indicator (LI) field contains the length of the information field. Thus, the total length of the frame is $LI + header\ size + offset\ size + FCS\ size$. The header CRC field is applied only to the LI field. The size and content of the offset field can be negotiated at set-up time. It can be used to transmit control messages from the transmitter to the receiver, or to identify the protocol used in the payload, or to indicate the QOS required by the payload, or to carry an MPLS label. The FCS field is calculated using all the bytes of the payload field including the offset field and the information field. The size of FCS field can be negotiated at set-up time. When IP packets are carried over SDL, the information field contains a single IP packet.

Length Indicator 2 bytes	Header CRC 2 bytes	Offset 0-32 bytes	Information 0-64 KB	FCS 0-4 bytes
-----------------------------	-----------------------	----------------------	------------------------	------------------

Figure 5: The SDL frame

2.3.2 Encapsulation of SDL frames into SONET

SDL over SONET works in a similar way as PPP over SONET. SDL frames are first scrambled, then placed row by row and byte-aligned within the payload of a SONET frames. One SDL frame may lie between two successive SONET frames. During the time that there are no IP packets available for transmission, the SONET payload is filled with idle SDL frames. An idle SDL frame is just a 4-byte header consisting of the LI and header CRC fields.

Delineation of frames within the SONET payload is achieved by a state machine, which has three states, namely, the hunt state, the pre-sync state, and the sync state. The state machine starts in the hunt state. In this state, the frame is processed byte by byte in order to calculate the header CRC for the assumed header field. If and only if a correct header CRC is found, then a header is assumed to be found and the delineation process enters the presync state. At the presync state, the header CRCs are sought based on the length field. If a wrong header CRC is found, the delineation process goes back to the hunt state. If several consecutive correct header CRCs are found, the delineation process enters the sync state. At the sync state, the delineation process goes back to the hunt state if a wrong header CRC is found. This delineation process is very similar to the delineation process of ATM cells. The main difference is that when the process for delineating SDL frames is in the sync state, a wrong header CRC will cause the delineation process to go back to the hunt state. However, in the delineation process of ATM cells, several continuous wrong HECs are necessary before the delineation process to go back to the hunt state. The reason is that the SDL frame has a variable length while the ATM cell has a fixed size.

2.4 Conclusions

As mentioned above, in the IP/ATM/SONET scheme, IP packets need to be segmented in order to be transported over ATM. This is not needed in other two schemes. In [2], it was estimated that an extra 10% of the raw bandwidth will be used because of the ATM cell header. Furthermore, if we consider the AAL5 overhead and the distribution of IP packet size in the internet backbone, the average ATM overhead was estimated to reach to about 25% of the average IP packet size. Under the same distribution of IP packet sizes, the average PPP/HDLC overhead is only about 2% of the average IP packet size. As for SDL, this should have an even lesser overhead than PPP/HDLC. Thus, when the bandwidth is expensive, IP/PPP/HDLC/SONET and IP/SDL/SONET are more attractive than IP/ATM/SONET.

The delineation of IP/ATM/SONET and IP/SDL/SONET are all length-based, while the delineation of IP/HDLC/SONET is flag-based. The length-based delineation can produce constant transmission overhead, while the flag-based delineation cannot.

Finally, we note that PPP/HDLC is a data link layer protocol and provides a point-to-point connectivity. On the other hand, ATM is a packet-switched network that can

provide a number of services to IP-based applications, such as end-to-end connectivity, quality of service, and network management.

The main features of these three schemes are summarized in Table 1.

	ATM	PPP/HDLC	SDL
Segmentation and reassembling	Yes	No	No
Protocol overhead	Large	Small	Small
Delineation	Length-Based	Flag-Based	Length-Based
Data link layer protocol	No	Yes	Yes

Table 1: A Comparison of the three schemes for IP over SONET

3. IP over WDM

WDM stands for *wavelength division multiplexing*. It is a multiplexing technique that exploits the wavelength domain. As mentioned in the introduction, several different technologies have been developed or are being developed, that permits the transfer of data directly over WDM, such as wavelength-routed networks, optical packet switching, and optical burst switching.

A wavelength-routed network provides end-to-end permanent or switched connections, known as lightpaths. These connections may go over several hops, and on each hop the same wavelength is allocated. Different wavelengths can be allocated to different hops for the same connection, but this requires the use of converters. Currently, lightpaths are permanent connections, and they are set-up administratively using network management procedures. Two IP routers can be connected by a lightpath, one for each direction. IP packets can be transmitted between the two routers over these lightpaths.

In this manner, a wavelength-routed network may be used to interconnect a set of backbone IP routers. Multi-Protocol Lambda Switching [3], a technique that extends MPLS traffic engineering control to work with the optical cross-connects (OXC) in the wavelength routed network, may be used to establish lightpaths between an ingress and an egress IP router. With MPLS, an OXC provisions lightpaths by establishing a relation between an (input port, input wavelength) tuple and an (output port, output wavelength) tuple.

Deploying OXC with MPLS traffic engineering in the Internet backbone is a natural first step towards an optical Internet. This approach employs WDM within the backbone, where it is mostly needed to alleviate packet processing and forwarding, while the rest of the Internet's current architecture remains as it is today. Specifically, IP packets are forwarded by conventional IP routers until they reach an ingress backbone router. Based on their destination, IP packets are sent by the ingress IP router on an appropriate wavelength so that they travel to the egress IP router through the backbone of OXC completely within the optical domain. Beyond the egress router, the IP packets are again forwarded by conventional IP routers until they reach their final destination.

A next step towards an optical Internet is to replace conventional IP routers by optical packet switches, in which packets remain in the optical domain while being switched

from input to output port. A further step would be to implement a different switching technique, namely optical burst switching, in addition to the traditional circuit switching and packet switching techniques.

In the remaining sections of this paper, we will review optical packet switching and optical burst switching. These techniques have been designed to transport IP traffic and ATM traffic in a more efficient way than wavelength-routed networks. The reader is referred to [16] for more information of the joint design of the electronic and optical layer. We note that in describing the optical devices below, we will make use of various optical components. These components are described in the appendix.

4. WDM optical packet switches

A WDM optical packet network consists of optical packet switches interconnected by fiber links. A WDM optical packet switch switches incoming optical packets to their correct output ports. As will be seen, the switching of the packets is done in the optical domain. Two switches are interconnected by fiber, running n different wavelengths. Also, it is also possible that two switches be interconnected by lightpaths.

Because of synchronization requirements, optical packet switches are typically designed for fixed-size packets (cells). In order for an IP packet to be carried in a network of optical packet switches, it must first be segmented into fixed-size cells, and the initial IP packet must be reassembled at the destination. The first study to provide insight into this process of segmentation and reassembly can be found in [15].

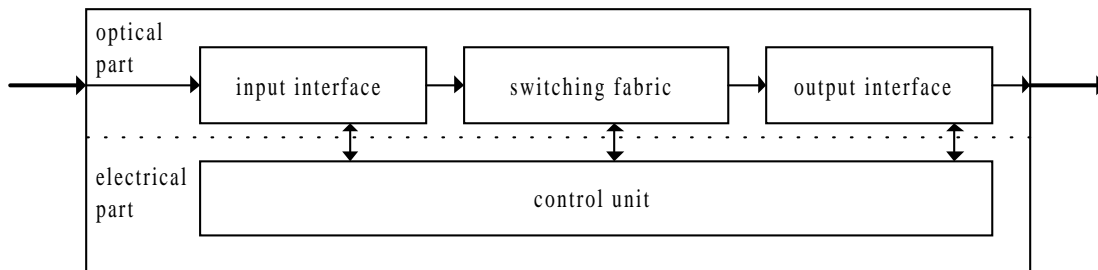


Figure 6: The WDM optical packet switch architecture

4.1 The optical node

The generic architecture of a WDM optical packet switch is shown in Figure 6. It consists of four parts, namely, the input interface, the switching fabric, the output interface, and the control unit. The input interface is mainly used for packet delineation and alignment, packet header information extraction and packet header removal. The switch fabric is the core of the switch and it is used for switching packets optically. The output interface is used to regenerate the optical signals and insert the packet header. The control unit controls the switch using the information in the packet headers.

When a packet arrives at a WDM optical packet switch, it is first processed by the input interface. The header and the payload of the packet are separated, and the header

is processed by the control unit electronically. The payload remains as an optical signal throughout the switch. After the payload passes through the switching fabric, it is combined with the header at the output interface.

In the following, we briefly describe some issues of optical packet switches. For more information about synchronization and contention resolution, the reader is referred to [23].

4.1.1 Packet coding techniques

Several optical packet coding techniques have been studied [5]. There are three basic categories, bit-serial, out-of-band-signaling, and bit-parallel. Bit-serial coding has been demonstrated using optical code division multiplexing, optical pulse interval and mixed rate techniques. In optical pulse interval coding, the packet header is processed optically. In mixed rate coding, the packet header is processed electronically at a slower rate, while the payload is processed optically at a higher rate. Out-of-band-signaling coding includes sub-carrier multiplexing (SCM) and dual wavelength coding. In SCM, the packet header is placed in an electrical subcarrier above the baseband frequencies occupied by the packet payload, and both are transmitted in the same time slot. In dual wavelength coding, the packet header and payload are transmitted in separate wavelengths but in the same time slot. In bit-parallel coding, multi-bits are transmitted at the same time but with separate frequencies.

4.1.2 Contention resolution

Contention resolution is necessary in order to handle the situation when more than one packet are destined to go out of the same output port at the same time. This is a problem that commonly arises in packet switches, and it is known as external blocking. It is typically resolved by buffering all the contending packets, except one which is permitted to go out. In optical packet switches, there are three ways to realize the contention resolution, discussed below.

Optical buffering

In the optical domain, buffers are implemented by *optical delay lines* (ODL). An ODL can delay a packet for a specified amount of time, which is related to the length of the delay line. There are many ways to use ODLs to emulate electrical buffers. For instance, a buffer for N packets with a FIFO discipline can be implemented using N delay lines of different length. Delay line i can delay a packet for i timeslots. There is a counter that keeps track of the number of the packets in the buffer. It is decreased by 1 when a packet leaves the buffer, and it is increased by 1 when a packet enters the buffer. Suppose that the value of the counter is j when a packet arrives at the buffer, then the packet will be routed to the j -th delay line. Limited by the length of the delay lines, this type of buffer usually has small size. For more information, the reader is referred to [12].

Exploiting the wavelength domain

In WDM, several wavelengths run on the fiber link that connects two optical switches. This can be exploited to minimize external blocking as follows. Let us assume that

two packets are destined to go out of the same output port at the same time. Then, they can be still transmitted out but on two different wavelengths. This method can be further improved by using optical buffering for each wavelength.

Deflection routing

Deflection routing is ideally suited to switches that have little buffer space. When there is a conflict between two packets, one will be routed to the correct output port, and the other will be routed to any other available output port. In this way, no or little buffer is needed. However, the deflected packet may end up following a longer path to its destination.

Below, we examine various optical packet switch architectures that have been proposed in the literature. Based on the structure of the switching fabric used, these switches have been classified to the following three classes: *space switch*, *broadcast-and-select*, and *wavelength routing*. For presentation purposes, we do not show the input and output interfaces and the control unit used in these switches. A description of the various optical components used in the optical devices presented below is given in the Appendix.

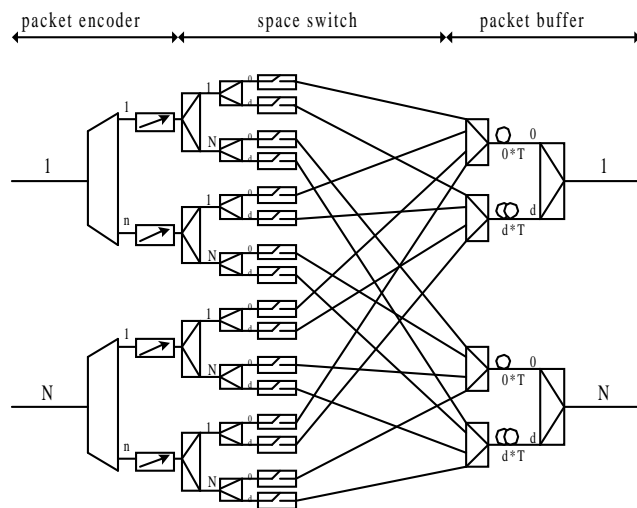


Figure 7: An architecture with a space switch fabric

4.2 An architecture with a space switch fabric

A space switch fabric architecture is shown in Figure 7. The performance of this switch was analyzed in [9]. The switch consists of N incoming and N outgoing fiber links, with n wavelengths running on each fiber link. The switch is slotted, and the length of the slot is such that an optical packet can be transmitted and propagated from an input port to an output optical buffer.

The switch fabric consists of three parts: optical packet encoder, space switch, and optical packet buffer. The optical packet encoder works as follows. For each incoming fiber link, there is an optical demultiplexer which divides the incoming optical signal

to the n different wavelengths. Each wavelength is fed to a different tunable wavelength converter (TWC) which converts the wavelength of the optical packet to a wavelength that is free at the destination optical output buffer. Then, through the space switch fabric, the optical packet can be switched to any of the N output optical buffers. Specifically, the output of a TWC is fed to a splitter which distributes the same signal to N different output fibers, one per output buffer. The signal on each of these output fibers goes through another splitter which distributes it to $d+1$ different output fibers, and each output fiber is connected through an optical gate to one of the ODLs of the destination output buffer. At the beginning of the slot, an optical packet from a wavelength is sent to its TWC, where it is converted to a wavelength that is free at the destination output buffer. Then, the optical packet is forwarded to an ODL by appropriately keeping one optical gate open, and closing the remaining. The information regarding which wavelength a TWC should convert the wavelength of an incoming packet and the decision as to which ODL of the destination output buffer the packet will be switched to is provided by the control unit, which has knowledge of the state of the entire switch.

Each output buffer is an optical buffer implemented as follows. It consists of $d+1$ ODLs, numbered from 0 to d . ODL i delays an optical packet for a fixed delay equal to i slots. ODL 0 provides zero delay, and a packet arriving at this ODL is simply transmitted out of the output port. Each ODL can delay optical packets on each of the n wavelengths. For instance, at the beginning of a slot ODL 1 can accept n optical packets, one per wavelength, and delay them for 1 slot. ODL 2, can accept n optical packets at the beginning of each time slot, and delay them for 2 slots. That is, at slot t , it can accept n packets (one per wavelength) and delay them for 2 slots, in which case, these packets will exit at the beginning of slot $t+2$. However, at the beginning of slot $t+1$, it can also accept another batch of n optical packets. Thus, a maximum of $2n$ packets may be in transit within ODL 2. Similarly for ODL 3 through d . Let c_i denote the number of optical packets on wavelength λ_i , where $i=1,2,\dots,n$. We note that these c_i optical packets may be on a number of different ODLs. To insert an optical packet into the buffer, we first check all the c_i counters to find the smallest one, say c_s , then we set the TWC associated with this optical packet to convert the packet's wavelength to λ_s , increase c_s by one, and switch the optical packet to the c_s ODL. If the smallest counter c_s is larger than d , the packet will be dropped.

4.3 Architectures with broadcast-and-select switch fabric

In this section, we describe two different architectures with a broadcast-and-select switch fabric. A broadcast-and-select architecture is such that packets from all input ports, each on a different wavelength, are combined within the switch and are broadcast to all the output ports. Wavelength selectors are then used at each output port to select a wavelength, and consequently, a packet, to be sent out the switch. This type of switch fabric lends itself to multicasting.

4.3.1 The KEOPS broadcast-and-select switch fabric

This switch [18][11][8][14][19] was developed as part of the European ACTS KEOPS (Keys to Optical Switching) project. Each input and output fiber carries only one wavelength, as shown in Figure 8. However the wavelength of an output port is not fixed, and varies with packets. The output interface is responsible for making it meet

the requirement of the output signal. The switching fabric consists of three blocks, namely the encoder, the buffer block, and the selector block. The wavelength encoder block consists of N fixed wavelength converters, one per input. The buffer block consists of K ODLs and a space-switching stage implemented by means of splitters and optical gates. Finally, the wavelength selector block consists of N wavelength channel selectors implemented by means of demultiplexers and optical gates. These three blocks make up the broadcast-and-select switch fabric.

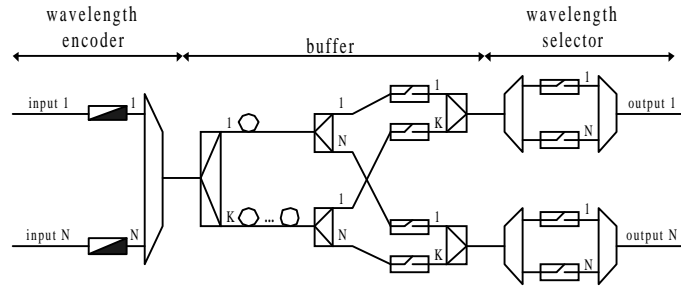


Figure 8: The KEOPS broadcast-and-select switch fabric

The switch is slotted. At the beginning of a time slot, each wavelength converter in the wavelength encoder block converts the wavelength of the incoming packet to a fixed wavelength. The output of the N converters is combined and then distributed through a splitter into K different ODLs. Each ODL has a different delay which is an integer number of slots. That is, ODL i has a delay of i slots. The N optical packets are stored simultaneously to the K different ODLs. At the beginning of the next slot, a maximum of KN optical packets exit from the K ODLs and up to N of them are directed to their destination output ports without any collisions. This is achieved through a combination of splitters, optical gates, demultiplexers and multiplexers. Specifically, the output signal from each ODL goes through a splitter which distributes it over N outputs. We recall that this output signal consists of N multiplexed optical packets, one for each wavelength. The signal from output j of each splitter is directed to output port j . Since there are K splitters, there are K such output signals, of which only one is selected and directed to output port j . This selected output signal is fed into a demultiplexer, which breaks it up to the N wavelengths, of which only one is transmitted out. The operation of this broadcast-and-select switch fabric is managed by a control unit.

In this switch, an optical packet consists of a header, a payload, and a guard time. The header may include information about the destination, payload type, priority, and so on. The payload is the user data. A guard time is used to allow for the set-up time of the optical devices in the switch. It may be inserted between the header and the payload, or between two successive packets on the same wavelength. A mixed rate coding is used. That is, the header is encoded at a low fixed bit rate (e.g. 622 Mb/s), and the payload rate may vary from a few hundred Mb/s to 10Gb/s. However, the packet length is fixed in time, and not in the number of bits. That is, the duration of the packet is fixed (e.g. 1.64 μ sec), but the size of the packet is variable. This packet format has the following two advantages. First, the processing speed of the logic in the WDM packet switches depends on the header rate, but does not depend on the payload rate. Secondly, the buffering space in the WDM packet switches, realized by means of ODLs whose length is proportional to the time length of the packet to be stored, does not depend on the payload rate.

This switch architecture can be extended to the case where it has M input and output fibers, and each input and output fiber carries N wavelengths. This is achieved by demultiplexing the signal from each incoming fiber to the N wavelengths, and then treating the switch as if it has NM input wavelengths instead of N presented above. At the output side, each group of N wavelengths can be combined together through a combiner onto the same output fiber. More information on this type of switch architecture can be found in [19].

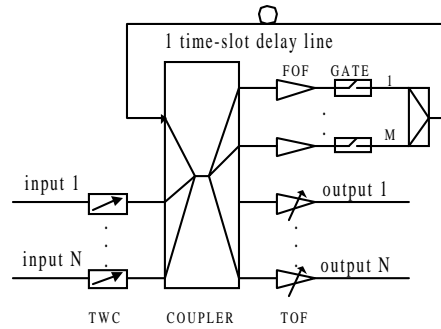


Figure 9: Broadcast-and-select with re-circulation buffer

4.3.2 Broadcast-and-select switch fabric with re-circulation buffer

This switch architecture was proposed in [4] and a modified version is shown in Figure 9. The idea of using a re-circulation buffer comes from an ATM switch, known as the *starlite* switch. As in the previous switch architecture, there is a single wavelength for each input and output fiber, and the wavelength of an output port varies with packets. The broadcast-and-select switch fabric is implemented through a coupler which combines up to M input wavelengths and then distributes the combined signal to N tunable optical filters and M fixed optical filters. Note that M is large than N . The input to the coupler comes from N input wavelengths and M wavelengths which are part of the feedback process, explained below.

The switch is slotted, and it is controlled by a control unit. At the beginning of each time slot, the control unit knows the destination output ports of the incoming optical packets from the input ports and the 1 time-slot delay line. Accordingly, it instructs the tunable wavelength converters at the input ports, the tunable optical filters at the output ports, and optical gates. Up to M optical packets are fed into the coupler, and according to their destinations, up to N of them are passed through the tunable optical filters and out to the output ports, and the remaining packets are re-circulated through an ODL. The re-circulated optical packets are fed back to the coupler at the beginning of the following slot.

4.4 Architectures with wavelength routing switch fabrics

In this section, we describe three switch fabrics based on wavelength routing switch fabrics. Briefly, the switching procedure of all these switches can be divided into two phases. In the first phase, packets are sent to ODLs for contention resolution, and in the

second phase, packets are routed to the correct output ports through the wavelength routing switch fabric.

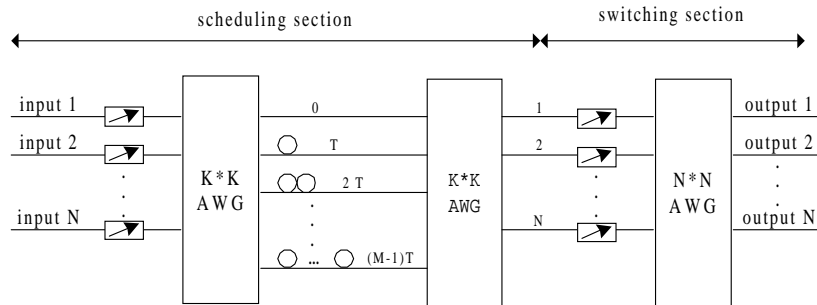


Figure 10: An input-buffered switch with a wavelength-routing fabric

4.4.1 An input-buffered wavelength routing switch fabric

This switch was proposed in [25], and it is shown in Figure 10. Each incoming and outgoing link carries a single wavelength. The wavelength of an output port varies with packets. The switch consists of the scheduling section and the switching section. The scheduling section is used for contention resolution and it is composed of N TWCs, one for each incoming wavelength, two $K \times K$ AWGs and M ODLs, where $K = \max(N, M)$. The combination of all these optical devices provides for an optical buffering of N individual buffers, each of them has M positions. If there are available buffer spaces, a packet entering input i of the first AWG will appear at output i of the second AWG after a specified delay. The length of the delay is determined by the wavelength of the packet when it enters the first AWG. Specifically, each TWC converts the wavelength of an incoming optical packet, so that the optical packet when routed through the AWG joins the ODL with the appropriate delay. The delay of an optical packet is selected using the following two rules: first, no two optical packets may appear at the same slot at the same switch output, and, secondly, no two optical packets may appear at the same buffer output at the same slot.

The switching section is used for switching optical packets to their destination output ports and is made up of an AWG and TWCs. The TWCs are used to assign the optical packet the right wavelength corresponding to the desired output port.

The switch suffers from the head-of-line blocking which is inherent in input buffering switches. For example, suppose that optical packet 1 in input i must be routed to output 1, while optical packet 2 behind optical packet 1 in input i must be routed to output 2. If optical packet 1 must be delayed for 1 time slot, optical packet 2 will have to be delayed for at least 1 time slot due to above rule 2, even though optical packet 2 goes to a different output port. However, in the case of optical packet 1 has to be delayed by more than 1 time slot, optical packet 2 need only be delayed by 0 slot as long as packet 2 has no conflict at the switch output 2.

In [25] the authors also proposed an output-buffering switch with a wavelength routing fabric.

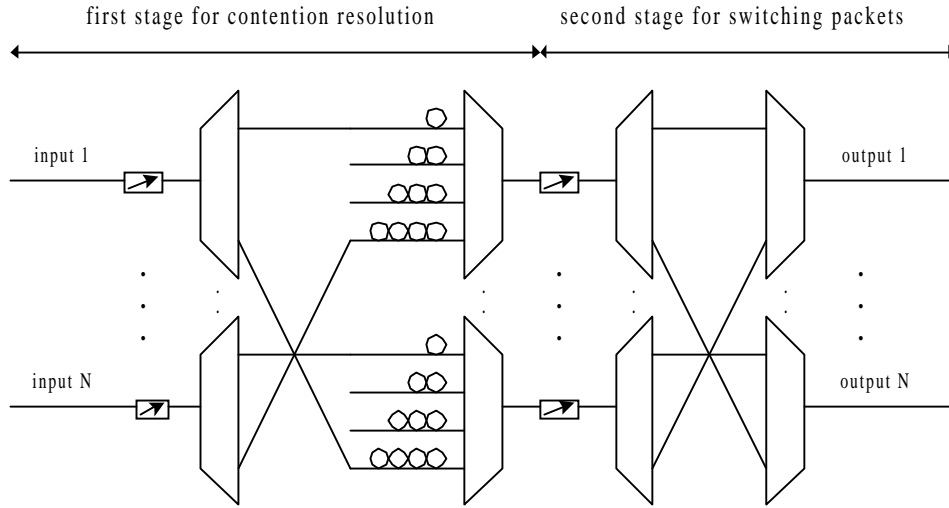


Figure 11: An input-buffered wavelength routing fabric with a distribution network

4.4.2 An input-buffered wavelength routing fabric with a distribution network

This switch was developed as part of the KEOPS project [18] [11], and it is shown in Figure 11. Each incoming and outgoing fiber carries a single wavelength. The wavelength of an output port varies with packets. The switch consists of two stages, namely the contention resolution stage and the switching stage. In the first stage, through the demultiplexer, each input port is connected with at least one ODL in each of N ODL sets. The TWC in the first stage decides to which ODL a optical packet will be sent. The second stage is used for switching optical packets to the correct output ports. Through the demultiplexer in the second stage, each ODL is connected with each output port. The TWC decides to which output port a optical packet will be sent.

Logically the first stage can be divided into two parts, a distribution part and an input buffer part. The distribution part distributes optical packets from the same input to different input buffers. Note that if we remove the distribution part, that is, each demultiplexer can only connect to one ODL set, the switch becomes identical to the one described above in section 4.4.1. The distribution part helps to overcome the head-of-line blocking. Let us refer to the example of the two optical packets discussed in the previous subsection. In this case, regardless of how many time slots optical packet 1 must be delayed, it has no effect on the delay of optical packet 2. That is because the two optical packets are routed to different ODL sets.

If an optical packet arrives at the switch at time slot t , it is routed to the delay line with length d determined by the following three constraints. First, no optical packet is scheduled to the same output port in the time slot $t+d$. Second, no optical packet is scheduled to the same TWC in the second stage in the time slot $t+d$. Third, no optical packet from the same input port and to the same output port was scheduled in d' with $d' \geq d$. The ODL with the shortest delay satisfying these three constraints is selected.

In the WDM version of this switch, there are M incoming and outgoing fiber links, and

each carries N wavelengths. N planes are located between M demultiplexers and M combiners connected to the incoming and outgoing fibers, respectively. Each plane is an $N \times N$ standard wavelength routing switch, as described above. However, in the WDM switch, an outgoing fiber may carry more than one wavelength. The control part of the switch makes sure that there is no wavelength conflict in the combiner.

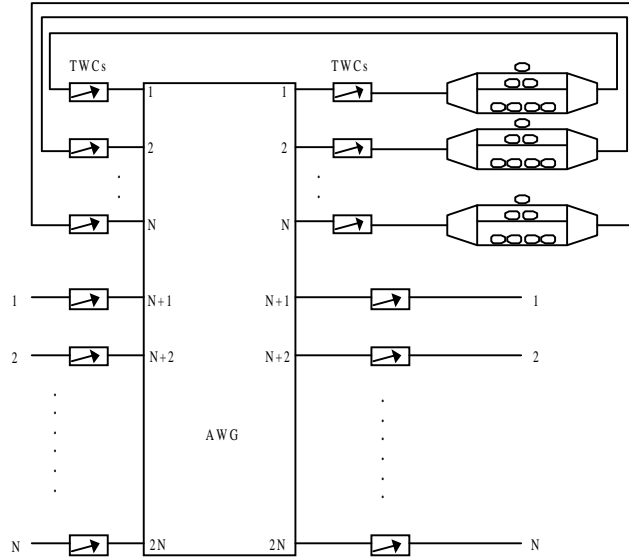


Figure 12: The WASPNET switch

4.4.3 A wavelength routing switch fabric with a re-circulation buffer

This switch was proposed as part of the WASPNET (wavelength switch optical packet network) project [13]. The configuration of a WASPNET switch with single-wavelength inputs and outputs is shown in Figure 12. It consists of a $2N \times 2N$ AWG, N sets of ODLs and $4N$ TWCs. As in the previous two wavelength routing switch fabrics, the switch can be divided into two phases. First, optical packets are routed to the ODLs to resolve contention, then they are routed to the desired output port. However, in this switch, these two phases are implemented together by a $2N \times 2N$ AWG and N ODL sets. The $2N$ TWCs on the left of the AWG are used to select the AWG's output. The first N TWCs on the right of the AWG are used to select the correct ODLs for the optical packets that will be re-circulated. The other N TWCs are used to convert optical packets to the wavelengths required by the switch output interface, because there are more wavelengths inside the switch than the incoming and outgoing wavelengths. One advantage of this switch is that it can support optical packet priorities. That is, after leaving the delay line, an optical packet maybe delayed again because of preemption by a higher-priority optical packet.

The WDM version of this switch is made up of demultiplexers, combiners, and multiple planes of wavelength routing switch fabrics with the single-wavelength inputs and outputs, as shown in Figure 13. It has N inputs and N outputs, each of them with n wavelengths. There are n planes, each corresponding to one of the n wavelengths. For example, wavelength i in each input fiber is always demultiplexed to plane i . In view

of this, the inputs of each plane have the same wavelength. However, different wavelengths may appear at the output of each plane. In one time slot, the switch allows multiple optical packets to leave not only from the same output of the WDM switch, but also from the same output of a single plane. The $N \times N$ AWG is used for this function. The N TWCs ($N+1, \dots, 2N$) that connect to the $2N \times 2N$ AWG carry out more functions, such as they make a final routing decision in addition to assigning optical packets to the wavelengths required by the switch output interface. The control part of the switch makes sure that there is no wavelength conflict in the combiner.

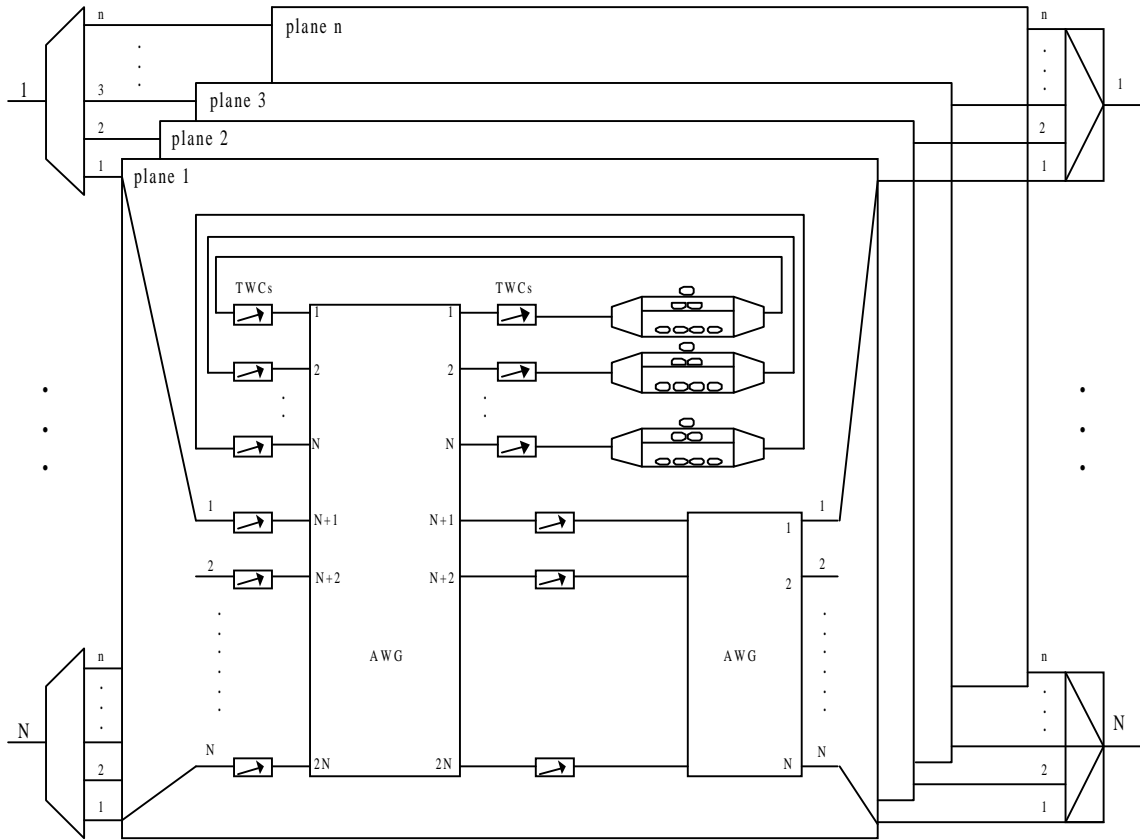


Figure 13: The WDM version of the WASPNET switch

5. Optical burst switching

Optical burst switching (OBS) is an adaptation of an ITU-T standard for burst switching for ATM networks, known as *ATM block transfer* (ABT). There are two versions of ABT, namely, *ABT with delayed transmission* and *ABT with immediate transmission*. In the first case, when a source wants to transmit a burst, it sends a packet to the ATM switches which are on the path of the connection to inform them that it wants to transmit a burst. If all the switches can accommodate the burst, the request is accepted and the source is allowed to go ahead with its transmission. Otherwise, the request is refused and the source has to send another request later on. In *ABT with immediate transfer*, the source sends the request packet, and then immediately after it, without receiving a confirmation, transmits its burst. If a switch along the path cannot carry the burst, due to congestion, the burst is dropped.

These two techniques have been adopted for optical networks. The *tell-and-go* (TAG) scheme [21], [22] is similar to the ABT with immediate transmission, and the *tell-and-wait* (TAW) scheme [22] is similar to ABT with delayed transmission. An intermediate scheme, known as *just-enough-time* (JET) was proposed in [17].

An OBS network consists of optical burst switches interconnected with WDM links. An optical burst switch transfers a burst coming in from an input port to its destination output port. Depending upon the switch architecture, it may or may not be equipped with optical buffering. The transmission links in the OBS network can carry multiple channels, and an optical burst is assigned to a channel dynamically. The channels are implemented using WDM or *optical time division multiplexing* (OTDM). The control packet associated with a burst may be also transmitted over one of these channels, or it may be transmitted over a non-optical network. The length of a burst may vary from one to several IP packets. Currently, OBS networks do not exist. An optical burst switch architecture is described in [20] and [7]. In these papers, various algorithms for scheduling bursts within an optical burst switch are also discussed. Also, an optical IP router adopting OBS is described in [6].

In the tell-and-go scheme, the source transmits the control packet and immediately after it transmits the optical burst. In this scheme, it may be necessary to buffer the burst in the optical burst switch, until its control packet has been processed. In the JET scheme, there is a delay between the transmission of the control packet and the transmission of the optical burst. This delay can be set to be larger than the total processing time of the control packet along the path. This way, when the burst arrives at each intermediate node, the control packet has been processed and a channel on the output port has been allocated. Therefore, there is no need to buffer the burst at the node. This is a very important feature of the JET scheme, since optical buffers are difficult to implement. A further improvement of the JET scheme can be obtained by reserving resources at the optical burst switch from the time the burst arrives at the switch, rather than from the time that its control packet is processed at the switch.

In [24] a variation of JET was proposed, which supports quality of service. Specifically, two traffic classes were defined, namely, real-time and non-real-time. A burst belonging to the real-time class is allocated a higher priority than a burst belonging to the non-real-time class, by simply using an additional delay between the transmission of the control packet and the transmission of the burst. The effect of this additional delay is that it reduces the blocking probability of the real-time burst at the optical burst switch.

6. Conclusions

In this paper, we have reviewed the different approaches to carrying IP traffic over optical networks, including IP over SONET, optical packet switching, and optical burst switching. The different schemes for transferring IP packets over SONET address the pressing problem of integrating the Internet with the widely deployed SONET infrastructure. The obvious next step is to develop solutions for exploiting the vast new bandwidth and flexibility available with (dense) WDM technologies. Already, there is significant work in progress in two areas: the evolution of the SONET infrastructure to include WDM, and the design and evaluation of IP over WDM solutions based on

Multi-Protocol Lambda Switching. Both areas are promising, with practical results and deployment expected very soon. On the other hand, optical packet switching and optical burst switching techniques are currently at an experimental stage, and there is definitely a need for new, high-performance and cost-effective architectures. As optical device mature and become cost-competitive, it is conceivable that by the end of the decade optics will play a much more significant role in switching technology than it does today, with many switching and routing functions incorporated into the optical domain.

References

- [1] B. S. Arnaud, "Architectural and engineering issues for building an optical internet", In Proceedings of SPIE, pages 358-377, SPIE, November 1998.
- [2] J. Anderson et al., "Protocols and architectures for IP optical networking", Bell Labs Technical Journal, pages 105-124, January-March 1999.
- [3] D. O. Awduche, Y. Rechter, J. Drake, R. Coltun, "Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control with Optical Crossconnects". draft-awduche-mpls-te-optical-00.txt
- [4] G. Bendeli et al., "Performance assessment of a photonic ATM switch based on a wavelength controlled fiber loop buffer", OFC'96 Technical Digest, pages 106-107, OFC, 1996.
- [5] D. J. Blumenthal, P. R. Prucnal, and J. R. Sauer, "Photonic packet switches: Architectures and experimental implementations", Proceedings of the IEEE, 82(11):1650-1667, November 1994.
- [6] F. Callegati, H. C. Cankaya, Y. Xiong, and M. Vandenhoute, "Design issues of optical IP routers for Internet backbone applications", IEEE Comm. Magazine, pages 124-128, December 1999.
- [7] Y. Chen and J. S. Turner, "WDM burst switching for petabit capacity routers", Proceeding of Milcom 1999 (IEEE Press)
- [8] D. Chiaroni, "Optical node for fast packet switching networks in the KEOPS project: Structure and performance aspects", In Proceedings of SPIE, pages 315-327. SPIE, November 1998.
- [9] S.L. Danielsen et al., "Analysis of a WDM packet switch with improved performance under bursty traffic conditions due to tunable wavelength converters", Journal of Lightwave Technology, 16(5):729-735, May 1998.
- [10] B. T. Doshi et al., "A simple data link protocol for high-speed packet networks", Bell Labs Technical Journal, pages 85-103, January-March 1999.
- [11] C. Guillemot et al., "Transparent optical packet switching: The european ACTS KEOPS project approach", Journal of Lightwave Technology, 16(12):2117-2133, December 1998.
- [12] D. K. Hunter, M. C. Chia, and I. Andonovic, "Buffering in optical packet switches", Journal of Lightwave Technology, 16(12):2081-2094, December 1998.
- [13] D. K. Hunter et al., "WASPNET: A wavelength switched packet network", IEEE Communications Magazine, pages 120-129, March 1999.
- [14] F. Masetti et al, "Design and performance of a broadcast-and-select photonic packet switching architecture", 22nd European Conference on Optical Communication ECOC'96, pages 309-312, ECOC, September 1996.
- [15] M. W. McKinnon, H. G. Perros, G. N. Rouskas, "Performance Analysis of Broadcast WDM Networks Under IP Traffic". Performance Evaluation, vol. 36-37, pp. 333-358, August 1999.
- [16] E. Modiano, "WDM-Based Packet Networks", IEEE Communications Magazine, pp. 130-135, March 1999.
- [17] C. Qiao and M. Yoo, "Optical burst switching (OBS) - A new paradigm for an optical internet", Journal of High Speed Networks, 8:69-84, 1999.

- [18] M. Renaud et al., "Network and system concepts for optical packet switching", IEEE Communications Magazine, pages 96-102, April 1997.
- [19] L. S. Tamil et al., "Analysis of all-Optical IP routers", Proceedings of SPIE, pages 456-464. SPIE, November 1998.
- [20] J. Turner, "Terabit burst switching", Journal of High Speed Networks, 8: 3-16, 1999.
- [21] E. Varvarigos and V. Sharma, "The ready-to-go virtual-circuit protocol: a loss-free protocol for multigigabit networks using FIFO buffers", IEEE/ACM Trans. Networking, 5: 705-718, October 1997.
- [22] I. Widjaja, "Performance analysis of burst admission-control protocols", IEEE proceedings communications, 142:7-14, February 1995.
- [23] S. Yao, B. Mukherjer, and S. Dixit, "Advances in Photonic Packet Switching: An Overview", IEEE Communications Magazine, pp. 84-94, February 2000.
- [24] M. Yoo and C. Qiao, "A new optical burst switching protocol for supporting quality of service", Proceedings of SPIE, pages 396-405. SPIE, November 1998.
- [25] W. D. Zhong and R. S. Tucker, "Wavelength routing-based photonic packet buffers and their applications in photonic packet switching systems", Journal of Lightwave Technology, 16(10):1737- 1745, October 1998.

Appendix: WDM Components

The various optical components used in the optical devices discussed in this paper are listed in Table 2.













Optical Delay Line	
Optical Gate	
Combiner	
Splitter	
Coupler	
Fixed Wavelength Converter (FWC)	
Tunable Wavelength Converter (TWC)	
Fixed Optical Filter (FOF)	
Tunable Optical Filter (TOF)	
Multiplexer	
Demultiplexer	
Arrayed-Waveguide Grating (AWG)	

Table 2: WDM Components

An *optical delay line* (ODL) is an optical fiber used to store optical signals for a specified duration. ODL technology is quite different from the electronic random-access memory. First, one can not access it at any time as in random access memory. For example, if at time t_1 , a packet is sent into an ODL, then at time t_2 , the packet will appear at the other end of the ODL. The difference between t_1 and t_2 is a function of the length of the ODL. Because of limitations on the length of ODLs, it is difficult to implement large buffers using ODLs.

An *optical gate* is used to either let an optical signal through or stop it.

A *combiner* combines two or more optical signals, each coming in from a different input optical fiber, into a single output optical fiber.

A *splitter* distributes the same optical signal from an input optical fiber to a number of output optical fiber. It is the reverse of a combiner.

A *coupler*, in general, is a combiner followed by a splitter.

A *fixed wavelength converter* (FWC) converts the optical signal from an incoming wavelength to another outgoing wavelength. Note that, the outgoing wavelength is fixed, whereas the incoming wavelength may not be fixed. There are two kinds of FWCs, namely, fixed-input and fixed-output converters, and variable-input and fixed-output converters.

A *tunable wavelength converter* (TWC) converts the optical signal from an incoming wavelength into any outgoing wavelength. There are two kinds of TWCs, namely, fixed-input and variable-output converters, and variable-input and variable-output converters.

A *fixed optical filter* (FOF) is an optical device that allows only one fixed wavelength

to pass through.

A *tunable optical filter* (TOF) is an optical device that allows any selected wavelength to pass through. TOFs are mainly characterized by their tuning range and tuning time.

A *multiplexer* combines incoming optical signals of different wavelengths, each coming in from a different input port, onto a common output port. The difference between a multiplexer and a combiner is that only one specified wavelength is allowed on each input port of a multiplexer, while there is no requirement for the type and the number of wavelengths on each input port of a combiner.

A *demultiplexer* divides an optical signal coming in from its input port, to different individual wavelengths which are transmitted out from different output ports. It is the reverse of a multiplexer. The difference between a demultiplexer and a splitter is that different wavelengths appear at different output ports of a demultiplexer, while all wavelengths appear at all output ports of a splitter.

An *arrayed-waveguide grating* (AWG) is a wavelength-routing device that can route optical signals from different input ports to different output ports based on their wavelength. Fixed routing is used. Figure 14 shows another wavelength routing device with fixed routing, which has the same routing as an AWG.

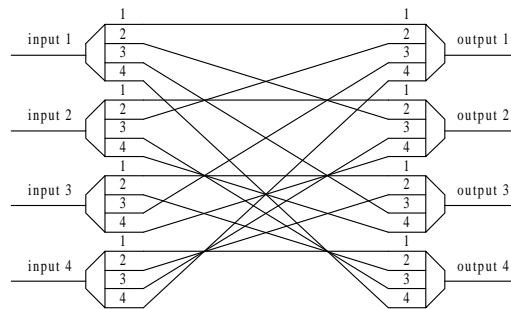


Figure 14: Fixed wavelength-routing