An Optical Packet Switch Based on WDM Technologies

X. Yu\textsuperscript{(a)}, J. Lin\textsuperscript{(a)}, X. Zhao\textsuperscript{(a)}, J. P. Zhang\textsuperscript{(a)}, Y. Gu\textsuperscript{(a)}, F. S. Choa\textsuperscript{(a)}, Guansong Zhang\textsuperscript{(b)}, Longjun Li\textsuperscript{(b)}, Huiping Xiang\textsuperscript{(b)}, Haldun Hadimioglu\textsuperscript{(c)} and H. Jonathan Chao\textsuperscript{(b)}

Abstract

Dense Wavelength-division multiplexing (DWDM) technology offers tremendous transmission capacity in optical fiber communications. However, switching and routing capacity lags behind the transmission capacity, since most of today’s packet switches and routers are implemented using slower electronic components. Optical packet switches are one of the potential candidates to improve switching capacity to be comparable with optical transmission capacity. In this paper, we present an optically transparent ATM (OPATM) switch that consists of a photonic front-end processor and a WDM switching fabric. A WDM loop memory is deployed as a multi-ported shared-memory in the switching fabric. The photonic front-end processor performs the cell delineation, VPI/VCI overwriting, and cell synchronization functions in the optical domain under the control of electronic signals. The WDM switching fabric stores and forwards aligned cells from each input port to the appropriate output ports under the control of an electronic route controller. We have demonstrated with experiments the functions and capabilities of the front-end processor and the switching fabric at the header-processing rate of 2.5Gb/s. Other than ATM, the switching architecture can be easily modified to apply to other types of fixed-length payload formats with different bit rates. Using this kind of photonic switches to route information, an optical network has the advantages of bit rate, wavelength, and signal-format transparencies. Within the transparency distance, the network is capable of handling a widely heterogeneous mix of traffic, including even analog signals.

Index Terms

\textsuperscript{(a) Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Baltimore, MD 21250}
\textsuperscript{(b) Department of Electrical and Computer Engineering, Polytechnic University, Brooklyn, NY 11201}
\textsuperscript{(c) Department of Computer and Information Science, Polytechnic University, Brooklyn, NY 11201}
Optical packet switch, WDM loop memory, switching fabric, cell delineation, VPI/VCI overwrite, cell synchronization.
An Optical Packet Switch Based on WDM Technologies

I. INTRODUCTION

The fast evolution of optical technologies, such as optical add-drop multiplexing (OADM) [1], [2], reconfigurable photonic switching [3], and dense wavelength-division multiplexing (DWDM) [4], have not only provided tremendous transmission capacity, but have also created a paradigm shift for the next generation network. For instance, the success of early multi-wavelength optical network (MONET) trials [5] has stimulated network providers to widely deploy WDM networks to meet the exponential growth of Internet traffic, where a number of optical cross-connect switches (OXC) [6] are interconnected by DWDM channels and are dynamically configured by network management.

Given the rapid increase in network traffic level and transmission bandwidth, switching nodes are required to have matching capacities. Optical cross-connect switches, such as the Lucent LambdaRouter [7], are usually configured in a larger time scale, e.g., a few tens of minutes, a few hours, days, or even longer. In addition, the switching granularity of OXCs is still at the SONET/SDH circuit level or at the wavelength (\(\lambda\)) level. Because of the fluctuation of IP traffic and the need to meet different quality of service (QoS) requirements for various kinds of traffic, IP routers and ATM switches are required to perform forwarding (switching) on a per packet or cell basis. However, the capacity of today’s ATM switches and IP routers is far behind optical transmission capacity due to the complexity of required processing, the large number of interconnections and bandwidth limitations of electronics components. For instance, the capacity of a core router from Juniper [8] in its basic configuration is only 640Gb/s, compared with the terabits per second capacity of a single WDM fiber. Thus, building a large-capacity packet switch using only electronic technology can potentially lead to a system bottleneck when interconnecting a number of electronic devices or modules into a large system. This capacity mismatch between optical transportation and electronic switching nodes results in a speed gap and calls for optical packet switch platforms that can provide a very large switch capacity with the granularity at the packet or cell level. Furthermore, optical packet switching has another
advantage which is the adaptation of the IP packet or ATM cell onto the WDM layer directly to reduce the overhead of traditional IP (ATM) / SONET (SDH) / WDM adaptation.

All-optical switching is the key to the success of the next-generation optical network due to its transparency, high speed and large capacity [9]. To flexibly use the potential huge capacity of optical fiber, several all-optical switching paradigms have been proposed and under intensive study. Here, we briefly compare several switch architectures and the corresponding scheduling algorithms in the time-slotted all-optical switching schemes, such as optical packet switching (OPS) [10]–[13], time sliced optical burst switching (TSOBS) [14], and optical cell switching (OCS) [15]. In these schemes, time is divided into slots of fixed size, and each timeslot is referred to as an optical cell, or a cell. Note that the terms ”packet” and ”cell” are used synonymously in this paper. Cells that enter an optical switch must be aligned with respect to their slot boundaries. Unaligned cells at an input port can be aligned with an optical synchronizer (OSYN). One implementation of OSYN will be presented later in this paper. However, since cells arrive at different inputs of the optical switch in an uncoordinated fashion, they may desire for the same output port in the same timeslot. Therefore, fiber delay lines (FDLs) are needed to delay (buffer) cells when contention occurs. The architecture of optical-buffered switch and the corresponding scheduling algorithms are the most challenging issues to the time-slotted all-optical switching schemes.

To buffer cells, optical timeslot interchangers (OTSI) have been widely employed. An OTSI is a single-input optical device which consists of a number of FDLs. Let $T_{cell}$ be the length of each timeslot and $F \cdot T_{cell}$ the maximum delay that can be imposed on a cell. Fig. 1(a) and 1(b) depict a nonblocking OTSI and a blocking OTSI, respectively. An OTSI is said to be nonblocking if it can rearrange any positions of cells without blocking as long as there is no timeslot conflict. If in some cases blocking may occur in the OTSI even there is no timeslot conflict, then the OTSI is said to be blocking. The implementation complexity of the nonblocking OTSI is very high, thus in practice, the blocking OTSI is a more attractive solution to performing timeslot interchange.
With reference to Fig. 2, a feedback-buffered optical-packet switch based on the arrayed waveguide gratings (AWG) devices has been proposed in [10] for OPS. In this switch architecture, the switching plane is combined with $N$ OTSIs, which are placed on the output side and able to feed the delayed packets back to the input side of the switch. In the center, an AWG switch fabric is employed. Each inlet of the AWG switch fabric is associated with a tunable wavelength converter (TWC). The TWCs are needed because AWG devices switch optical signals according to their wavelengths. For those packets that have lost in the contention, they are assigned delay values and switched to the proper OTSIs for buffering. Such a buffered switch has been proven to be able to provide a low loss rate and low average delay for OPS. However, the scheduling algorithms that can efficiently assign delay routes for optical packets by using blocking OTSIs have not received enough attention in [10].

Time sliced optical burst switching (TSOBS) is a variant of optical burst switching (OBS) [14], in which burst contention is resolved in the time domain rather than the wavelength domain, thus eliminating the necessity for wavelength conversion that occurs in the traditional OBS schemes. In [14], the authors have also proposed an efficient scheduling algorithm for the per-input-OTS1 optical switch. The architecture of the per-input-OTS1 optical switch is given in Fig. 3, where the OTSIs are the blocking ones as shown in Fig. 1(b). In this algorithm, the existing schedule (switch configuration) is formulated as a directed graph, which gives all possible delay paths for data bursts. The assignment problem can thus be formulated as a searching problem in the directed graph. Nevertheless, since the algorithm assumes an OTSI is employed at each input port, the implementation complexity of the switch is undesirably high.

In [12], Karol has proposed a single-stage shared-FDL switch for optical packet (ATM) switch. The structure of the single-stage shared-FDL switch is given in Figure 4. The switch contains a number of feedback FDLs that are shared among all input ports. Suppose that there are $Z$ feedback FDLs, $N$ input ports, and $N$ output ports. Each FDL can delay cells by a fixed number of timeslots, and any two FDLs may have the same or different delay values. The outputs (inputs) of FDLs and the inputs (outputs) of the switch are collectively called the inlets (outlets) of the
switch fabric, yielding $N+Z$ inlets and $N+Z$ outlets. Karol has also proposed a non-reservation scheduling algorithm for the switch in which he assumed that the delay values of the $Z$ FDLs are all different from $1$ to $T_{cell}$ to $ZT_{cell}$. The algorithm is said to be non-reservation because there is no reservation (hence no departure time scheduling) for the cells that have lost in the contention and need to be buffered. That is, in each timeslot, cells can only be matched with the output ports for the current timeslot. For those buffered cells, there is no guarantee that they can get access to the desired output ports after coming out from the FDLs. Therefore, they may need to face another round of contention. Minimum reservation can be achieved by giving the highest priority to the oldest cell when resolving contention. However, since departure time is not scheduled in advance, the delay bound of Karol’s algorithm can be very large and it may require a cell to be switched and circulated many times. This is undesirable because the optical signals get attenuated each time when they are switched. Another issue of Karol’s algorithm is the high time complexity, which is $O(Z^2)$, due to its sequential nature.

Optical cell switching (OCS) is an optical timeslot wavelength switching scheme [15]. By using time-division multiplexing (TDM), time in OCS is divided into time slots of fixed size, and every $F$ consecutive time slots constitute a frame. The switching configuration of each core switch is changeable from slot to slot, and is repeated from frame to frame. Wavelengths in each time slot are switched as a bundle. While OCS can perform traffic grooming, and support both guaranteed and best-effort services the slot alignment remains to be a challenging issue. Table I compares time-slotted all-optical switching paradigms. The OPATM, which is an OPS, has the same OPS features, except that its scheduling is simpler and out-of-sequence packets do not occur.

[Table 1 about here.]

In this paper, we explore switching ATM cells in the optical domain by prototyping an OPATM switch based on WDM technology [16]. In the OPATM prototype, ATM cells are transferred back to back over the WDM layer directly without any adaptation layer in between, since ATM by itself has link layer functions. By taking the advantage of both optical and electronic technologies, we route ATM cells through an optical switching plane, while extracting and processing their headers in an electronic plane that, in turn, controls the optical devices and routes the cells to the proper output port(s). Although today’s optical packet switching technology is still very
primitive and cannot compete with electronic switching technology, optical packet switches have
great potential to scale up their switching capacity as the technology of some key optical devices
becomes mature. This paper addresses such key devices and presents our results toward making
the optical packet switches as competitive as electronic switches.

Comparing with the aforementioned architectures, the advantages of this design are its sim-
plicity of the cell scheduler and elimination of cell out of sequence occurrences inside the
switch. The other architectures above may have out of sequence cells since the cells can be
stored in different fibers. Thus, the scheduler needs to keep track of the location of every stored
cell associated with each connection (or flow) which increases its complexity. The processing
time is so constraint that the cell size has to be increased to a frame, made of several packets,
increasing the delay through the optical switch node. More detailed comparisons regarding to
the implementation will be discussed in section IV.

The OPA TM switch consists of a photonic front-end processor and a switch fabric. In the
photonic front-end processor, there are three main functional units, cell delineation, VPI/VCI
overwrite, and cell synchronization. The cell delineation unit identifies cell boundaries. By not
transmitting cells in SONET frames, the possibility of having variable gaps between or within
cells caused by the need to carry SONET overhead bytes is eliminated. We adopt the ITU-T
standard to perform cell delineation by finding the correct cyclic redundancy check (CRC) at
the fifth byte position of the ATM cell header. A similar method has been also been proposed
to identify packet boundaries for IP over WDM [17]. Once a cell boundary is identified, the
VPI/VCI field is replaced optically with the new value by the VPI/VCI overwrite unit. The cell
synchronization unit is used to align the phases of incoming ATM cells in the optical domain.
The synchronization issue is also addressed in [18]. In our design, we adjust the phases of optical
cells at 2.5Gb/s to a reference cell clock with the adjustment range from 1 to 511 bits and with
a precision of $\frac{1}{2}$ bit (or 100ps). We adopt a novel sampling method to achieve 100ps precision
without using a 10GHz clock. Finally, the switch fabric of the OPATM switch is an optical
shared memory based on the WDM technology [19]. It uses the broadcast and select method
to achieve the switching and multicasting functions by employing wavelength converters and
tunable filters at the input and output ports.

Using OPATM switches to route information, an optically transparent network has the advant-
tages of having bit rate, wavelength, and signal format transparencies. The signal formats can be
IP, ATM, FDDI, etc. for digital signals and AM, FSK, QAM, DQPSK, etc. for analog signals. Within the transmission distance (transparency distance) that a signal can have a considerable S/N ratio, an OPATM network can be transparent to all types of signals. The advantage is obvious: one network can transport all different types of traffics.

The rest of the paper is organized as follows: Section II describes the architecture of the OPATM switch, including the photonic front-end processor and the optical switch fabric. We describe the implementation and testing of the OPATM switch and the experimental results obtained from the testing in section III. In section IV we discuss the advantages and disadvantages of our implementation when compared with other works. We also discuss the scalability of implementation. Finally, section V presents the conclusions.

II. ARCHITECTURE OF THE OPATM SWITCH

Fig. 5 shows the architecture of an $N \times N$ OPATM switch, where incoming cells are split into two paths. Cells on the top path remain in the optical domain and are routed through the optical switch plane. Cells on the bottom path are converted to the electronic domain, where their headers are extracted for processing (e.g. finding the output ports for which the cells are destined and finding new VPI/VCI values to replace the old ones). The electronic central controller performs cell delineation, VPI/VCI overwrite, cell synchronization, and routing. The first three functions are implemented in the photonic front-end processor, and the last one is handled by the route controller that routes cells to their destination output port(s).

As shown in Fig. 6, the cell format adopted in our system has 64 bytes with 5 bytes of header, 48 bytes of payload, and 2 guard time fields (with all ones), which are 6 and 5 bytes long. The guard times are used to accommodate the switching times of optical devices, such as optical tunable filters. The guard field lengths are arbitrarily chosen to demonstrate the feasibility of the technology.

The incoming optical cells in Fig. 5, after being delayed by fiber lines, processed for their headers, and synchronized by the cell synchronization unit, are sent to the switch fabric. In the
switch fabric, cells are converted to different wavelengths by wavelength converters (WCs) that are controlled by the route controller, which keeps track of the available wavelengths in the WDM optical shared memory. Cells read from the WDM optical shared memory are broadcast to all $N$ ports by a $1 \times N$ splitter and selected by the destination output port (or ports, if multicast) through tunable filters that are tuned by the route controller at a per-cell basis. The final wavelength converter stage converts cells to their predetermined wavelengths. Since the broadcast-and-select switching mechanism is the underlying mechanism, implementing multicasting is straightforward. The front-end processor presented in this paper is used with a header line rate of 2.5Gb/s, although the payload can be at any desired bit rate.

We introduce in the following subsections the structure of the front end processor and the switch fabric. A preliminary study of optical packet switches is presented in [20] that paved the way for the current work. Details of the front-end processor unit are given in [21] and [22].

A. The Front End Processor

Fig. 7 shows the front-end processor. An optical cell stream is tapped from each input line, converted to electronic format, and sent to the cell delineation unit. Cell delineation is a process used to identify cell boundaries so that the incoming cell stream can be further processed at the cell level by other units, such as the VPI/VCI overwrite unit. We adopt the standardized header error code (HEC) checking mechanism to find cell boundaries. It takes the advantage of the inherent CRC coding correlation between the cell header to be protected (the first four bytes) and the HEC byte (the fifth byte of the cell header).

Identifying cell boundaries for a back-to-back cell stream at 2.5Gb/s is very challenging. In our design, when 32 mismatched HEC checking are detected in the HUNT state, the input bit stream is shifted by one bit to restart the HEC hunting process [21]. This is achieved by masking a clock pulse at 2.5GHz (or 400ps). The 16-bit parallel CRC circuit, as shown in Fig. 8, includes 58 XOR gates organized in four levels and this circuit works at a clock frequency of 155MHz (or 6.4ns cycle time).
Once cell boundaries are recognized and confirmed the state machine enables the VPI/VCI overwrite unit with the cell clock and signal X, as shown in Fig. 9. The main function of this unit is to overwrite the VPI/VCI field of the incoming cell header in the optical domain. The VPI/VCI overwrite unit performs table lookups in the electronic domain to retrieve the new VPI/VCI, then converts it to an optical signal, and replaces the old VPI/VCI with the new one. The challenge is how to handle the high-speed overwrite at the bit rate of 2.5Gb/s with each bit taking only 400ps. We resolve it by (a) replacing the whole cell header instead of just the VPI/VCI field, and (b) using electronic variable delay lines (programmable delay) to compensate for the time difference between the old header and the new header. The new header converted to serial and then used to control a laser driver to generate the new cell header in the optical domain, replaces the old one by a 2×1 optical switch. The successfully overwritten cells are sent to fiber delay lines in the cell synchronization unit.

The cell synchronization unit shown in Fig. 10 optically aligns cells from different inputs to the extent of $\frac{1}{4}$ bit (100ps or 2cm optical delay line at 2.5Gb/s) before entering the switch fabric. In the unit, we divide the control into two steps. A coarse adjustment circuit controls the first nine optical delay elements and adjusts the phases of incoming cells down to the bit level. A fine adjustment circuit controls the last two and further adjusts the phase down to $\frac{1}{4}$ bit.

Each stage of the optical delay element consists of a Y-junction SOA gate, a combiner, and a fiber delay line with a delay of $T/2^n$ (where $T$ is one cell time and $n$ is from 1 to 11). In order to address challenging issues in optical delay line fabrication, such as polarization, noise reduction, coherent crosstalk, and power stabilization, we have fabricated InP-based semiconductor Y-junction switches which are described below.

B. Optical Switch Fabric

All incoming cells, after processed by the photonic front-end processor, are buffered and forwarded to their destined output port(s) through a WDM loop memory. Fig. 11(a) shows the
architecture of the WDM loop memory, which has $3N$-port array waveguide grating (AWG) routers and with a common fiber delay line (one cell delay). Each cell uses a specific wavelength in the loop. Therefore the total capacity of this optical memory is $N$ cells, or equivalently, $N$ wavelengths permitted to exist in this memory at the same time. Fig. 11(b) shows the stored and read out signals after 15 and 24 circulations inside the loop memory. There are two SOA gates used as switches to control the cells remaining in the loop or going out of the loop. The control signals of SOA gates are from SOA controllers in the route controller. Fig. 12 shows a $4 \times 4$ switch fabric and its route controller. In the switch fabric, a coupler is used to combine new cells from the inputs and existing cells in the loop. An existing cell needs to be re-circulated in its loop if there is at least one copy of the cell that is not a head of line (HOL) cell in one of the output queues. Thus, this cell needs to be stored in its loop until all copies are read out.

![Figure 12 about here.]

The optical devices in the system need to have a flat gain over wide bandwidth to accommodate WDM applications. The gain of the optical switches used in our system is kept just enough to overcome the insertion loss while the input power is sufficiently high to suppress the amplified spontaneous emission (ASE) noise.

III. IMPLEMENTATION AND TESTING

In this section, we discuss the implementation and testing of the OPATM switch we have prototyped. We first describe the implementation of optical devices and electronic circuits. We then show experimental results and our current work whose results can aid in scaling up the switch for higher-speed and higher-capacity networking applications.

A. Fabricated optical devices

In the photonic front-end processor and the optical switch fabric, there are three key optical devices, SOA switches, wavelength converters, and tunable filters. They are described below.

1) SOA switches and their integration:

![Figure 13 about here.]
We have developed integrated $1 \times 2$ Y-junction SOA switches to use in optical delay elements of the cell synchronization unit and the WDM optical shared memory. Fig. 13 shows the integrated $1 \times 2$ Y-junction SOA photonic switch, which has a size of $2 \times 0.5 \text{mm}^2$. It is fabricated on an InP substrate by wet etching and regrowth processing. The Y-junction passive waveguide, with a stripe width of $3.75 \mu m$, has a buried rib structure with two output ports separated by $250 \mu m$. The total length of the passive waveguide is $1200 \mu m$. The SOA active regions are located at both the input and output sides of the device to overcome the insertion loss. Each one of the SOAs has a length of $400 \mu m$. They are made of InGaAs-InGaAsP multiple quantum wells around 1550nm. After the antireflection coating, the ripple in the spontaneous emission spectrum is smaller than $0.2 \text{dB}$ at the working bias current.

Fig. 14 shows switching characteristics of the fabricated SOA switches. An optical "dc" signal is switched on and off by the SOA gate controlled by external electrical signals. Both the rising and falling times are around 600ps, which are partially limited by the driving electronics. Further improvement of the speed by reducing the area of the electrical contact 5 to 10 times is possible. However, since guard times between cells are employed, the sub-nanosecond switching speed of the device is enough for our application. By using the fabricated $1 \times 2$ SOA switch, we have demonstrated data-block switching operations [23].

Fig. 15 shows an integrated 3-stage cell synchronizer and a bonded chip of the device based on the SOA switch basic building block. Underneath the pair of metal contacts are the SOA switches. The design requires three-side fiber coupling. Such kind of packaging is considerably difficult to implement. In the coupling process we have found that to couple light in and out of a semiconductor passive waveguide is extremely difficult. Adding active sections to the inputs and outputs of a chip is very helpful for the fiber coupling. Based on this experience, we have fabricated an improved version of the integrated delay line chip as shown in Fig. 16. The distance between the 2-output waveguides are exactly 250 $\mu m$. The active sections on each input and
output waveguide are 400 µm long. A multi-stage cell synchronizer can be built using such an array. The final size of the device can be much smaller. The fiber coupling can also be simplified from a 3-side fiber alignment to a 2-side fiber alignment.

2) Wavelength converters: In the OPATM switch, the wavelength converters are deployed in the switch fabric. Those in front of the optical shared-memory are used to convert the wavelengths of incoming cells to a number of idle wavelengths before being fed into the memory. The ones after the tunable filters are used to convert the wavelengths of the cells coming from the shared memory to the predetermined ones before they are transmitted out of the switch system.

![Figure 17 about here.](image1)

![Figure 18 about here.](image2)

![Figure 19 about here.](image3)

One can use the cross-gain and/or the cross-phase modulation effects [24] in SOA gates to do wavelength conversions. Our "gain decompression effect" work [25] has shown that we can use either the side injection [26] or the long SOA cavity length [27] to achieve high-speed wavelength conversions. We have made devices for implementing either scheme into our conversion devices. Fig. 17 shows pictures of the fabricated side-injected wavelength converters. In Fig. 17(a), an interaction amplifier integrated with a side-injection amplifier is shown. In 17(b), a 3-section tunable laser integrated with an interaction amplifier and a side-injection amplifier is shown. Fig. 18 shows pictures of two types of long-cavity λ converters: an all active Y interferometer and an all active Mach-Zender converter with separated injection branches. From our theoretical calculations and experiments we have found that the long-cavity SOA type of devices is more effective than the side-injected SOAs since the interaction time is longer. Using the long-cavity SOA devices, conversion results at 5 Gb/s are easily achieved as shown in Fig. 19.

3) Tunable filters:

![Figure 20 about here.](image4)

![Figure 21 about here.](image5)
Fast tunable filters are used in our work to select cells at each output port. We have tested our fabricated tunable active filters and used them as wavelength demultiplexers in an 8-wavelength system. It has an insertion gain instead of an insertion loss. Fig. 20 shows the experimental setup of using the active filter. The DBR filter is composed of gain, phase, grating, and post-filter gain sections. Error signals for wavelength control and dynamic gain control are extracted from the gain section and fed back to the grating section and the gain section, respectively. The post-filter gain section can be used to switch the output on and off. The central wavelength of the DBR filter can be easily tuned and locked to a desired channel by changing the grating bias current. As shown in Fig. 21, 8-channel WDM signals with 0.8nm (100GHz) spacing are coupled into the DBR gain section. A channel rejection ratio of 20 dB is achieved when the average input power of each channel is at -30dBm. The extinction ratio reduced to 15dB when the signal level is increased to -25dBm/channel, due to the gain saturation of the active filter. The saturation power of the device is improved by using a ridge type waveguide in our subsequent work.

B. Fabricated electronic circuit boards

The operations in the optical plane are under the control of signals coming from the electronic plane. In order to generate these signals, we have developed four kinds of printed circuit boards (PCBs). There are three kinds in the photonic front-end processor for the three functional units, cell delineation, VPI/VCI overwrite, and cell synchronization. The 4th kind of PCB is for the route controller.

The first three boards are implemented using off-the-shelf ECL and GaAs chips that can run at several gigabits per second. On the cell delineation PCB, a 16-bit demultiplexer performs serial-to-parallel conversions and converts the input bit stream to a 16-bit parallel format, reducing the number of ultra high-speed components. EPROM chips are used for VPI/VCI translation on the VPI/VCI overwrite PCB. A 16-bit multiplexer performs parallel to serial conversion for the cell headers and overwrites the old VPI/VCI optically. On the cell synchronization PCB, a substantial number of programmable delay chips are used to adjust the phase of the signals in the coarse and fine adjustment circuits. The delay adjustment range varies from 1390ps to 3630ps with approximate 20ps delay-step resolution. More detailed description of these three PCBs can be found in [20] and [21].
The primary component on the route controller PCB shown in Fig. 22 is a field programmable gate array (FPGA) chip that can be hardware programmed to perform the routing control functions. These include keeping track of the idle wavelengths, maintaining the output queues (FIFOs), and sending out appropriate control signals to the wavelength converters, SOA switches, and tunable filters in the switch fabric. The functions of the route controller can be even more complex. The fast re-programmability of the FPGA chip makes it flexible to accommodate various experiments on the optical switch fabric. The FPGA chip is an XCV300PQ240 FPGA which is a member of the Virtex FPGA family of Xilinx, Inc. Unlike the first three kinds of PCBs that are running at the speed from 155 Mb/s to 2.5Gb/s, the route controller operates at the cell level and the frequency of the main working clock is at 80MHz (16 times of the cell clock). The route controller PCB is designed as a 6-layer PCB.

The electrical power consumption of the boards are as follows: cell delineation board 40W, VPI/VCI Overwriting board 25W, synchronization board 41W and route controller 4W. Two cell delineation boards and two VCI Overwriting boards are used in our setup. The total power consumption is 175W for the electronic plane. The optical elements including lasers, modulators, switches, and receivers consume less than a 5 watts of total power. This does not include the power of the bias electronic and temperature control instruments, which are designed for general applications and would be unfair to include them into the picture. Although a great reduction of electrical power can be achieved by using higher density integrated electronics instead of multiple individual small scale ICs, the research still clearly shows that optical switching fabrics will be advantageous in both BW and power consumption compared with those of electronic switching fabrics.

C. Experimental results

We have built and demonstrated a 2×2 WDM ATM Multicast Switch using the devices and
circuit boards mentioned above. The detailed results are given in this section. Fig. 23 shows the overall setup. Fig. 24 shows the optical setup. Fig. 25 shows the FPGA route controller and the optical loop memory setup.

[Figure 26 about here.]

[Figure 27 about here.]

[Figure 28 about here.]

[Figure 29 about here.]

When the optical packets are sent into the two input ports, we use repeated patterns so that we can see them clearly on the oscilloscope. Fig. 26 shows the input packets. Two cell delineation units for the two input ports identify the input-cell boundaries, and generate synchronization signals, including cell clocks and byte clocks. As mentioned above, each cell delineation unit is placed on a PCB. Our experiments show that these two PCBs stably generate the clock signals and perform cell delineation at the speed of 2.1Gb/s and 2Gb/s. Our analysis indicates that this is due to the PCB placement arrangement of the chips that are responsible for the cell delineation operation. To be precise, the very high-speed CRC-based cell delineation circuit with GaAs and ECL chips are distributed all across the large PCB, causing large signal delays and noise pick up, and so the safe delineation range is below 2.5Gb/s. We have concluded that by modifying the PCB layout, the cell delineation would happen at 2.5Gb/s. We did not design the new cell delineation PCBs due the approaching end of the funded project. Overall, we performed all electronic experiments at 2.5Gb/s and at its derivatives, such as 155Mb/s, with the exception of the cell delineation unit.

The VPI/VCI overwrite units, as indicated above, replace old headers with new headers, while keeping the old payloads as shown in Fig. 27, 28 and 29. After the headers of input cells are replaced and before they are sent into the optical memory for contention resolution and route controlling, they have to be aligned in the time domain using the synchronization circuits. The synchronization signals including cell clocks, byte clocks and bit clocks are generated from the cell delineation units and sent to the synchronization circuit board to control variable optical delays. The same synchronization signals are used to synchronize route control signals and switch
packets in and out of the memory as well.

[Figure 30 about here.]

The cell synchronization PCB controls the 11 stages of optical delay elements and turns on or off the 12 Y-junction SOA switches. The fiber delay length varies from $\frac{1}{4}$, $\frac{1}{2}$, ..., to $\frac{1}{2n}$ of a cell time, where $n$ is the delay stage number. Fig. 30 demonstrates the $1/8$, $1/4$, and $1/2$ cell delays by using appropriate fiber length for an optical packet with 400 ns length at 2.5 Gb/s. The fiber-to-fiber insertion loss of the SOA switch is currently at 5 dB, achieving substantial contrast.

[Figure 31 about here.]

Since the major difficulty in aligning the phase is during the last few stages, here we describe the phase alignment of the last three stages of the cell synchronization unit. A reference clock is distributed in the whole switch system and used by all the inputs as a common alignment basis. The electronic part of the unit operating at 2.5 Gb/s and integrated with optical devices, generates optical delays as small as 100 ps. Fig. 31 show the delays generated from 100 ps to 400 ps. The accuracy is as small as a $\frac{1}{4}$ bit at 2.5 Gb/s.

[Figure 32 about here.]

[Figure 33 about here.]

In another experiment, we test the optical WDM loop memory controlled by an electronic route controller. The setup includes a WDM memory that has two input ports, two output ports and 2 sets of laser sources; each generating 4 different wavelengths (only two wavelengths for each port are used in this experiment). The testing setup is shown in Fig. 32. In the testing of the WDM memory fabric, we feed the two input ports with incoming cells and instruct the route controller which output ports these cells are destined. The destination port numbers are carried by the R1 and R2 signals as shown in Fig. 32 and 33. Since we are using only one pattern generator, the generated data (cells) is shared by both input ports. But each port may require its cell destined for a different output port. Thus, the route controller generates control signals W1-W4 and S1-S4 to accomplish the appropriate tasks requested by R1 and R2.
In the WDM memory setup of Fig. 32, the data coming from the pattern generator is modulated onto an optical signal with wavelength $\lambda_x$ and fed into the wavelength converter. The four laser sources in the bottom-left corner, L1-L4, provide the idle wavelengths to be converted to. The four signals coming from the route controller, W1-W4, control the four laser diodes and thus the allocated idle wavelengths. For example, W1 and W3 is logic '1' in the "Cell 1" slot means that $\lambda_1$ and $\lambda_3$ are allocated for "Cell 1" as shown in Fig. 29. (Note that for simplicity, we have decided here that $\lambda_1$ and $\lambda_2$ are used for the cells destined to output port 1 and $\lambda_3$ and $\lambda_4$ for output port 2 for ease of observations at the outputs.)

The wavelength-converted cells enter the loop memory through a 50:50 coupler. The erbium doped fiber amplifier (EDFA) here is used to compensate the power loss of the optical devices in the loop memory. The waveguide grating router (WGR) splits the multiplexed wavelengths into four separate paths, each passing a polarization controller (PC), a 1×2 SOA gate switch, and an attenuator. The four SOA switches are controlled by the route controller with S1-S4 signals to keep the cells in the loop memory (if the S signal is a logic '0') or read the cells out to an output port (if the S signal is a logic '1'). For instance, as shown in Fig. 33, "Cell 2" is converted to wavelength $\lambda_1$ in the second cell slot but it appears in the third cell slot of output 1 under the control of the logic '1' of S2 signal in the third cell slot. In fact, we apply the simple periodic data pattern of Fig. 33 to the testing since it covers most of the necessary devices to build a complete optical switch fabric and it demonstrates the most important properties of a packet switch, e.g., output contention resolution and packet buffering.

[Figure 34 about here.]

[Figure 35 about here.]

Fig. 34 and 35 show the switching control signals and test results photographed on the oscilloscope. The waveforms of "Output 1" and "Output 2" in Fig. 35 are the same as expected in Fig 33. Note that the time scale of the cells (A~F) is different from that of the outputs for the ease of illustration and only the headers and partial payloads of the cells are shown in the figure. In this experiment, with four wavelengths, it is either that a cell is immediately output or kept in the loop memory for a number of cell slots until it is its turn to be output. Note that increasing the number of wavelengths would reduce the cell loss as more cells can stay in the
loop memory. However, the contention at the output ports would rise and force some cells to stay in the memory for a longer duration.

[Figure 36 about here.]

Figure 36 shows the tested bit error rate for 1, 15, and 30 loops in the WDM memory. The bit rate is 2.488Gbit/s. One period of the input signal is composed of an ATM cell (64 bytes long) and 29 empty cells. A 2.488Gbit/s PIN receiver is used to receive the packets. The power penalty is measured to be approximately 1.5dB per 15 rounds. However, due to ASE noise accumulation, a noise floor is formed when the cells circulate more than 15 rounds. This noise floor limits the maximum cell storage time. To reduce the ASE noise, the total passive loss in the loop has to be minimized so a small gain EDFA can be used.

In [28] we show that cells stored in the WDM memory can be refreshed just like an electronic dynamic random access memory (DRAM). Furthermore, a detailed study on stabilizing the loop memory gain using the gain clamping technique is described in [29].

IV. DISCUSSIONS

In this section, we discuss the rationales of the approaches taken to implement the OPATM switch. We start with the front end processor. The way we swap the VCI (or label) is simple and straightforward. Although the subcarrier based approach [30], [31] has been frequently utilized, it is a more complex and expensive approach in terms of both implementation cost and bandwidth utilization. To synchronize the system, two sets of clock recovery systems are required for the header and payload, respectively. When the system is run at real packet mode, they may need two sets of frame boundary detection and synchronization subsytems, which can be very costly. The subcarrier has to be multi-GHz away from the base band signals, which limits the minimum channel spacing. To do label swapping requires not just to write in new header but also to clean up the old header. This will require the full optical packet to go through the subcarrier extraction processor [32], [33]. The signal distortion introduced in the process for either using a fiber grating [33] or a fiber loop mirror [32] is very detrimental and that can limit the number of all optical processors this optical signal can continue propagating through in the network. On the other hand, the simple header-switching approach used in this work introduces minimum effects to the optical signal itself and its followed transmissions.
For the switch fabrics, we use the share memory and WDM, broadcast-and-select (BAS) approach. Compared with using pure TDM approaches [34] our approach requires a much more relaxed speed performance for components used in the switch. This helps to reduce cost and complexity for future scaling up. On the other hand, the scalability of a pure TDM approach is very limited. The speed of optical components usually is limited by the material fundamental limits and cannot be dramatically increased.

Conventional space-division-multiplexing (SDM) approach is robust. However, the loss is increased in the order of $O(N^2)$ and cannot be scale up to large size easily unless optical amplifying devices like SOAs are integrated into the system to compensate the loss. The WDM BAS approach has the advantages of having a simplified switch fabric and the switch fabric insertion loss can be reduced to $O(N)$ [16]. However, to achieve such a performance, the switch requires a fast wavelength-tuning device. The fast tuning requirement is coming from the need of converting a cell wavelength to a new wavelength, which is currently not used in the shared memory. It is also required at the output port to convert cells to the desired wavelength before leaving the switch. The tuning time has to be shorter than the cell period, which is in the hundreds of nanoseconds range. Using a fast tunable semiconductor laser we can easily achieve such a tuning speed [35]. However, there are not many choices for fast tunable filters with a tuning speed. Most tunable filters are either slow or with a narrow tuning range. The below-threshold tunable laser approach we used in this work is a direct application of fast tunable lasers and the tuning speed is the same as a fast tunable lasers.

In terms of scalability, our recent work have been focusing on carrying heterogeneous traffic with different types of payloads and increasing the scalability of the switch fabric by improving the bandwidth of the optical devices. An all-optical packet network using digital packet headers and arbitrary payloads, including higher bit rate payloads (compared with that of the header) and analog payloads, is proposed and demonstrated in [36]. Ultra-broadband SOAs with a gain profile that allows more than thousands of ITU-grid wavelength channels have been developed. The SOA has a gain profile to cover a tunable laser tuning range from 1300 nm to 1550 nm [37]. Broadband wavelength conversions among all wavelengths inside such a broad range have also been demonstrated [38]. To improve the high noise and high crosstalk characteristics of SOAs as gain materials, new techniques such as quasi-indirect bandgap materials and type II MQW gain materials have also been developed [39], [40]. A large improvement of SOA crosstalk
characteristics using graded bandgap techniques has been demonstrated [41]. By implementing these new techniques and technologies, we can scale up the WDM BAS switch fabric with a size up to \( N > 1000 \) when the channel spacing inside the switch is 50 GHz (not required for external transmissions).

V. Conclusion

We present the architecture and implementation of an OPATM switch and the test results of the prototype. The OPATM switch embodies the complete solution to an optical switching system that includes both a photonic front-end processor and an optical switch fabric. On the incoming back-to-back ATM cells, the OPATM switch identifies cell boundaries following the ITU-T standards via CRC checking on the cell header, optically overwrites the VPI/VCI fields in the cell header, synchronizes all the incoming cells to a common system reference point, and then buffers and forwards the cells to their destination output ports. The cell synchronization unit is able to adjust the cell delay with a precision of \( \frac{1}{4} \) bit (or 100ps) without using a 10GHz clock. Novel optical devices and electronic circuits are fabricated during the prototyping process.

The OPATM is a time-slotted optical packet switch. Designed for ATM, the switching architecture presented here can be easily modified to accommodate other types of fixed-length payload formats with different bit rates. An optical network using OPATM type photonic switches has the advantages of bit rate, wavelength, and signal-format transparencies. Within the transparency distance, the network is capable of handling a widely heterogeneous mix of traffic, including even analog signals. Unlike other OPS systems, cell scheduling is simple and out-of-sequence cells do not occur. The switch fabric with the WDM optical shared memory uses the broadcast-and-select switching mechanism, resulting in a simple implementation of multicasting.

We have fabricated SOA switches, wavelength converters and tunable filters for the photonic front-end processor and the optical switch fabric. In addition, we have designed and fabricated six electronic boards, containing high-speed GaAs and ECL chips and a high-density FPGA chip. The FPGA chip has the advantages of programmability and high speed, which are required for fast prototyping and successful cell switching, respectively.

The demonstrated testing results prove the feasibility and correctness of our proposed schemes. However, as mentioned before, the optical switching technology is still in a relatively early stage and can be observed in our prototype and experimental process. Challenging issues would arise
especially when the number of ports of the switch system is increased to a large number. For instance, the synchronization of a large number of cells would place strain on the fine adjustment of cells in our current approach. Furthermore, the buffer requirement of today’s core routers is at the level of millions of cells, which far exceeds the capacity of the optical memory implemented here. We plan to pursue on these challenges in the future.

**REFERENCES**


## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Architectures of OTSIs</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>An optical buffered AWG packet switch</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>A per-input-OTS1 optical switch</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>A single-stage shared-FDL optical switch</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>The architecture of the OPATM switch</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>The OPATM adopted cell format</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>The Photonic Front End Processor</td>
<td>31</td>
</tr>
<tr>
<td>8</td>
<td>Block diagram of the cell delineation unit</td>
<td>32</td>
</tr>
<tr>
<td>9</td>
<td>Block diagram of the VPI/VCI overwrite unit</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>Block diagram of the cell synchronization unit</td>
<td>34</td>
</tr>
<tr>
<td>11</td>
<td>WDM loop memory</td>
<td>35</td>
</tr>
<tr>
<td>12</td>
<td>The 4x4 optical shared-memory switch fabric with route controller</td>
<td>36</td>
</tr>
<tr>
<td>13</td>
<td>A fabricated 1x2 SOA gate switch</td>
<td>37</td>
</tr>
<tr>
<td>14</td>
<td>Switching characteristics of an SOA switch (200ps/div)</td>
<td>38</td>
</tr>
<tr>
<td>15</td>
<td>Integrated Delay lines</td>
<td>39</td>
</tr>
<tr>
<td>16</td>
<td>The improved integrated fiber delay line chip</td>
<td>40</td>
</tr>
<tr>
<td>17</td>
<td>Fabricated side-injection wavelength converters</td>
<td>41</td>
</tr>
<tr>
<td>18</td>
<td>All active integrated wavelength converters</td>
<td>42</td>
</tr>
<tr>
<td>19</td>
<td>Wavelength Conversion to Different Wavelengths at 5Gb/s</td>
<td>43</td>
</tr>
<tr>
<td>20</td>
<td>Experimental set up for a fast active tunable filter</td>
<td>44</td>
</tr>
<tr>
<td>21</td>
<td>8-Wavelength channels: before and after the active filter</td>
<td>45</td>
</tr>
<tr>
<td>22</td>
<td>The PCB of the route controller</td>
<td>46</td>
</tr>
<tr>
<td>23</td>
<td>The experimental setup</td>
<td>47</td>
</tr>
<tr>
<td>24</td>
<td>The optical setup</td>
<td>48</td>
</tr>
<tr>
<td>25</td>
<td>Route controller and the loop memory</td>
<td>49</td>
</tr>
<tr>
<td>26</td>
<td>Two input packets</td>
<td>50</td>
</tr>
<tr>
<td>27</td>
<td>The old and new headers synchronized</td>
<td>51</td>
</tr>
<tr>
<td>28</td>
<td>New header and old payload</td>
<td>52</td>
</tr>
<tr>
<td>29</td>
<td>The old and new packets (headers swapped)</td>
<td>53</td>
</tr>
<tr>
<td>30</td>
<td>Coarse Delay Adjustment</td>
<td>54</td>
</tr>
<tr>
<td>31</td>
<td>Fine Delay Adjustments</td>
<td>55</td>
</tr>
<tr>
<td>32</td>
<td>Test setup for the optical loop memory</td>
<td>56</td>
</tr>
<tr>
<td>33</td>
<td>The timing diagram for the testing of the optical switch fabric</td>
<td>57</td>
</tr>
<tr>
<td>34</td>
<td>Switching control signals</td>
<td>58</td>
</tr>
<tr>
<td>35</td>
<td>Test results for the optical switch fabric</td>
<td>59</td>
</tr>
<tr>
<td>36</td>
<td>Bit Error Rate test results for the optical loop memory</td>
<td>60</td>
</tr>
</tbody>
</table>
Fig. 1. Architectures of OTSIs
TWC : Tunable wavelength converter
OTSI: Optical timeslot interchanger

Fig. 2. An optical buffered AWG packet switch
Fig. 3. A per-input-OTSI optical switch
Fig. 4. A single-stage shared-FDL optical switch
Fig. 5. The architecture of the OPATM switch
Fig. 6. The OPATM adopted cell format
Fig. 7. The Photonic Front End Processor
Fig. 8. Block diagram of the cell delineation unit
Fig. 9. Block diagram of the VPI/VCI overwrite unit
Fig. 10. Block diagram of the cell synchronization unit
Fig. 11. WDM loop memory

(a) Structure

(b) The stored and read out packet signals of the WDM memory
Fig. 12. The 4x4 optical shared-memory switch fabric with route controller
Fig. 13. A fabricated 1x2 SOA gate switch
Fig. 14. Switching characteristics of an SOA switch (200ps/div)
Fig. 15. Integrated Delay lines

(a) device chip

(b) bonded device
Fig. 16. The improved integrated fiber delay line chip
Fig. 17. Fabricated side-injection wavelength converters
Fig. 18. All active integrated wavelength converters
Fig. 19. Wavelength Conversion to Different Wavelengths at 5Gb/s
Fig. 20. Experimental set up for a fast active tunable filter
Fig. 21. 8-Wavelength channels: before and after the active filter
Fig. 22. The PCB of the route controller
Fig. 23. The experimental setup

The two VCI overwrite boards
The two cell delineation boards
The two synchronization control circuits on one board
Fig. 24. The optical setup
Fig. 25. Route controller and the loop memory
Fig. 26. Two input packets
Fig. 27. The old and new headers synchronized
Fig. 28. New header and old payload
Fig. 29. The old and new packets (headers swapped)
Fig. 30. Coarse Delay Adjustment
Fig. 31. Fine Delay Adjustments
Fig. 32. Test setup for the optical loop memory
**Fig. 33.** The timing diagram for the testing of the optical switch fabric

- **T1:** One cell slot (204.8ns at 2.5Gb/s)
- **T2:** Period of the repeated data pattern (5 cell slots)
Fig. 34. Switching control signals
Fig. 35. Test results for the optical switch fabric
Fig. 36. Bit Error Rate test results for the optical loop memory.
LIST OF TABLES

I  A comparison of time-slotted all optical switching paradigms          62
## Table I

A comparison of time-slotted all optical switching paradigms

<table>
<thead>
<tr>
<th>Optical switching paradigm</th>
<th>Bandwidth reservation</th>
<th>Loss rate</th>
<th>Optical buffer</th>
<th>Wavelength conversion</th>
<th>Other Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPS</td>
<td>Nil</td>
<td>Moderate</td>
<td>Required</td>
<td>Necessary</td>
<td>High scheduling complexity, packets out of order</td>
</tr>
<tr>
<td>OBS</td>
<td>Link-by-link</td>
<td>High</td>
<td>Not required</td>
<td>Necessary</td>
<td>Burst contention resolution in time domain</td>
</tr>
<tr>
<td>OTWS</td>
<td>End-to-end</td>
<td>Nil</td>
<td>Not required</td>
<td>Not necessary</td>
<td>Challenging slot alignment</td>
</tr>
</tbody>
</table>