ULTRA-LOW LATENCY MULTIPROTOCOL OPTICAL ROUTERS FOR THE NEXT GENERATION INTERNET

University of California

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2003-122 has been reviewed and is approved for publication.

APPROVED: PAUL SIERAK
Project Engineer

FOR THE DIRECTOR: WARREN H. DEBANY, JR., Technical Advisor
Information Grid Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE MAY 2003	3. REPORT TYPE AND DATES COVERED Final Jun 00 – Sep 02

4. TITLE AND SUBTITLE

ULTRA-LOW LATENCY MULTIPROTOCOL OPTICAL ROUTERS FOR THE NEXT GENERATION INTERNET

5. FUNDING NUMBERS
C   -   F30602-00-2-0543
PE   -   62301E
PR   -   K210
TA   -   21
WU   -   A1

6. AUTHOR(S)

S.J. Ben Yoo

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of California
Dept. of Electrical & Computer Engineering
One Shields Avenue
Davis CA 95616

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Advanced Research Projects Agency   AFRL/IFGA
3701 North Fairfax Drive   515 Brooks Road
Arlington VA 22203-1714   Rome NY 13441-4505

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

AFRL-IF-RS-TR-2003-122

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: Paul Sierak/IFGA/(315) 330-7346

12a. DISTRIBUTION / AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

Research conducted under this program demonstrated the first all-optical label switching router, the first all-optical label swapping with negative penalty, the first edge-router for optical-label switching, the first cascaded operation of optical-label switching routers, and the first field trial of optical label–switching routers. The router exhibited the following performance: subnanosecond switching speed; contention resolution that achieved a packet loss rate below 0.5% at a load of 0.6; arbitration and contention resolution with ultra low 250 nsec latency; and label switching IP over WDM with a packet loss rate <0.1%.

14. SUBJECT TERMS

Optical router, Optical label switching, All optical label switching router

15. NUMBER OF PAGES
114

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

List of Figures

Tables

Technical Status Report

1. SUMMARY

This final report summarizes the entire progress made during the 27 month contract period for the project titled Ultra-Low Latency Multi-Protocol Optical Routers for the Next Generation Internet. The progress made during this contract period fulfilled all the required deliverables and significantly advanced the optical router technology. In particular, this project demonstrated the first all-optical label switching router, the first all-optical label swapping with negative penalty, the first edge-router for optical-label switching, the first cascaded operation of optical-label switching routers, and the first field trial of optical-label switching routers. The optical router has achieved the following key performance milestones.

- Sub-nanosecond (~600 psec) switching speed of the optical router based on rapidly tunable wavelength converter,
- Innovative contention resolution with wavelength-time-space domain switching aided by electronic buffering to achieve packet loss rate below 0.5 % at load 0.6.
- FPGA based arbitration and contention resolution with ultralow 250 nsec latency,
- Combination of the edge router and the optical label switching router demonstrating the first cascaded all-optical label switching IP over WDM with packet loss rate below 0.1 %.

2. TASK LIST and DELIVERABLE PLAN OVERVIEW

The project is performed by interworking and collaboration of the following tasks as per the Amendment of June 20, 2000. Following this amendment, the project was descoped on March 6, 2002.

This descoped amendment reflects reduction of the funding level from the amended funding level of $860,000 over 36 months to $550,000 over 28 months. This descoping is in addition to the amendment of June 20, 2000 which reflects reduction of the funding level from proposed $ 1,174,520 over 3 years including option, to the amended funding level of $860,000 over 3 years.

The descoped amended Task List, Deliverables, Milestones, and Budget are shown below in this document.

List of Tasks as per F30602-00-2-0543 and Descoped List

As per agreement No. F30602-00-2-0543, the following had been the original task list as of June 20, 2000. Through the descoping, we propose to eliminate Task 3d and the entire Task 4. are eliminated.

Task 1: Network and Optical Router Architecture
    Task 1a: Network Architecture and Interoperability studies
    Task 1b: Optical Router Architecture Design
    Task 1c: Optical Switching and Contention Resolution tradeoff evaluation

Task 2: Optical Technologies
    Task 2a: Arrayed Waveguide Grating Design and Prototyping
    Task 2b: Rapidly Tunable Wavelength Converter Design and Prototyping
    Task 2c: Optical-Label Header Processor Design and Prototyping

Task 3: Systems Integration and NGI Supernet
    Task 3a: Switch Fabric Integration
    Task 3b: Controller and Driver Integration
    Task 3c: Header Processor Integration
    Task 3d: Optical Router and Testbed Integration

Task 4: NGI Supernet demonstration
> Task 4a: Systems Integration of additional Optical Routers
> Task 4b: Systems Integration of NGI Supernet and LAN
> Task 4c: Collaborative NGI applications and experimentation

Descoped List of project is as follows.

Task 1: Network and Optical Router Architecture
> Task 1a: Network Architecture and Interoperability studies
> Task 1b: Optical Router Architecture Design
> Task 1c: Optical Switching and Contention Resolution tradeoff evaluation

Task 2: Optical Technologies
> Task 2a: Arrayed Waveguide Grating Design and Prototyping
> Task 2b: Rapidly Tunable Wavelength Converter Design and Prototyping
> Task 2c: Optical-Label Header Processor Design and Prototyping

Task 3: Systems Integration and NGI Supernet
> Task 3a: Switch Fabric Integration
> Task 3b: Controller and Driver Integration
> Task 3c: Header Processor Integration

The following are **the descoped list of final deliverables**.

Task 1: Network and Optical Router Architecture
> Deliverable 1a: Optical Router Architecture Design Complete (Y1Q1)
> Deliverable 1b: Router Interoperability Study of Optical-Label Switching, Opical-MPLS, Terabit-Burst Switching, and Optical-Flow switching (Y1Q4)
> Deliverable 1c: Optical Switching and Contention Resolution tradeoff evaluation complete (Y1Q3)

Task 2: Optical Technologies
> Deliverable 2a: Optical-Label Header Processors Operational (Y2Q1)
> Deliverable 2b: Arrayed Waveguide Grating Operational (Y2Q3)
> Deliverable 2c: Rapidly Tunable Wavelength Converter Operational (Y2Q4)

Task 3: Systems Integration and NGI Supernet
> Deliverable 3a: Header Processor with Forwarding Table Integration Complete (Y2Q2)
> Deliverable 3b: Controller and Driver Integration Complete (Y3Q1)
> Deliverable 3c: Switch Fabric Integration Complete (Y3Q2)

3. PERSONNEL:

This project is conducted in *Optical Switching and Communications Systems Laboratory* at UC Davis. This group has been established in March of 1999, and currently involves the following researchers. Marked with asterick (*) are those who are directly supported by DARPA/AFRL Agreement No. F30602-00-2-0543.

Prof. S. J. Ben Yoo* : Principal Investigator (Task 1, 2, 3, 4)

Dr. Minyong Jeon* : Research Scientist (Task 2, 3, 4)
Yanda Zhang: Graduate Student (Task 2, 3, 4)
Srikanth Vaidianathan: Graduate Student (Task 2, 3, 4)
Yash Bansal: Graduate Student (Task 1, 3)
Charles Blas: Graduate Student (Task 1, 3)
Shun Yao: Graduate Student (Task 1, 3)
Steven Fong: Graduate Student (Task 2)
Vincent Tsui: Graduate Student (Task 2)
Vincent Hernandez: Graduate Student (Task 4)
Zhong Pan: Graduate Student (Task 1, 3)
Jing Cao: Graduate Student (Task 2)
Julie Taylor: Graduate Student (Task 1,3)
Yixue Du: Graduate Student (Task 2)
Zuqing Zhu: Graduate Student (Task 2, 3)
Zubin Wang: Graduate Student (Task 2, 3)
Danjue Li: Graduate Student (Task 1)
Haijun Yang: Graduate Student (Task 1, 3)
Dr. Fred Xue: Research Scientist (Task 1)
Dr. Il-Young Han: Research Scientist (Task 2)

In addition to the above UC Davis personnel, we are collaborating with the following researchers.

Dr. Steven Bonds (*Lawrence Livermore National Laboratory,* Task2, Material and Devices)
Dr. Sebastian Lourdoduss (*Royal Institute of Technology, Sweden,* Task2, Material and Devices)
Dr. Bill Lennon (*Lawrence Livermore National Laboratory,* Task 4, NTON Supernet)
Dr. Hal Edwards (*Nortel Networks,* Task 4, NTON Supernet)
Dr. Paul Daspit (*Nortel Networks,* Task 4, NTON Supernet)
Dr. Katsunari Okamoto (*NTT*, Task 2, Arrayed Waveguide Grating)
Dr. Kirk Boyer (*Tektronix,* Task 3, Universal Network Access System)
Prof. Joy Laskar[+] (*Georgia Tech*, Task 3, MMIC subcarrier circuit)
Dr. Rich Vohhannel[+] (*Corning*, Task 3, Optical Fiber transmission)
Profs. Jim Demmel[+], Jean Walrand[+] and 105 others[+] (*UC Berkeley and other UC campuses*, Task 4, CITRIS network for field trial).
Dr. Joachim Vobis[+] and Giovonnae Anderson[+] (*Agilent Technologies*, Task 4, CITRIS network for field trial).

Throughout the contract period, the PI has made important contacts and collaboration ties. The PI has established new collaboration relationships with Prof. Sebastian Lourdoduss in Royal Institute of Technology in Sweden, Prof. M. K. Smit in Technical University of Delft in Netherlands. Kirk Boyer of Tektronix have collaborated to pursue a universal network access module for the Optical label switching routers.

In addition, the PI's team was able to demonstrate the optical router in the field network through collaboration with NTON and Sprint. The field experiment has been carried out with half of the PI's group on the Livermore side, and the other half on the Sprint side. The experiment was successful, and marks the first field trial of Optical-Label switching and packet routing.

The PI's collaboration with 105 other UC faculty members bear a fruit this quarter. The California state's Science Initiative, Center for Information Technology for the Interest of Society (CITRIS) has been finally selected and its budget has been signed. In this proposal, dark fiber connection of four campuses,

Through collaboration and additional support from OIDA, the PI's group received two sets of 32x32 Arrayed Waveguide Grating Routers from Dr. Katsunari Okamoto.

4. FINAL REPORT PROGRESS SUMMARY (by Deliverables)

This section summarizes the content of each deliverable in a chronological order. The details for the deliverables are included in the previous quarterly reports.

4.1. Deliverable 1a: Optical Router Architecture Design Complete (Y1Q1):

The Optical Routers are enhanced with two types of label-processing modules to efficiently handle bursty traffic. Figure 1 shows a general schematic of an optical router. The first type of modules (LP-CI) interface between the client interfaces of the Optical Router and the client machines (*e.g.* IP routers) to encode optical signaling headers onto the packets added into the network, and to remove optical signaling headers from the packets dropping out of the network. The second type of modules (LP-TI) tap a small fraction of the optical signal from the Input Transport Interfaces (TI), detect signaling header information, and relay the appropriate commands to the switch fabric in the Network Element after looking up the forwarding table. The fiber delay is placed at the TI so that the packet including the header and the payload reaches the switch fabric after the switching occurs. This fiber delay will be specific to the delay associated with the combined time delay of header detection, table look-up, and switching. Our targeted goal is to accomplish this in ~100 nsec with a ~20 m fiber delay. Since there is no optical-to-electrical, electrical-to-optical conversion of the data payload at the Optical Router, the forwarding is independent of data format and protocol.
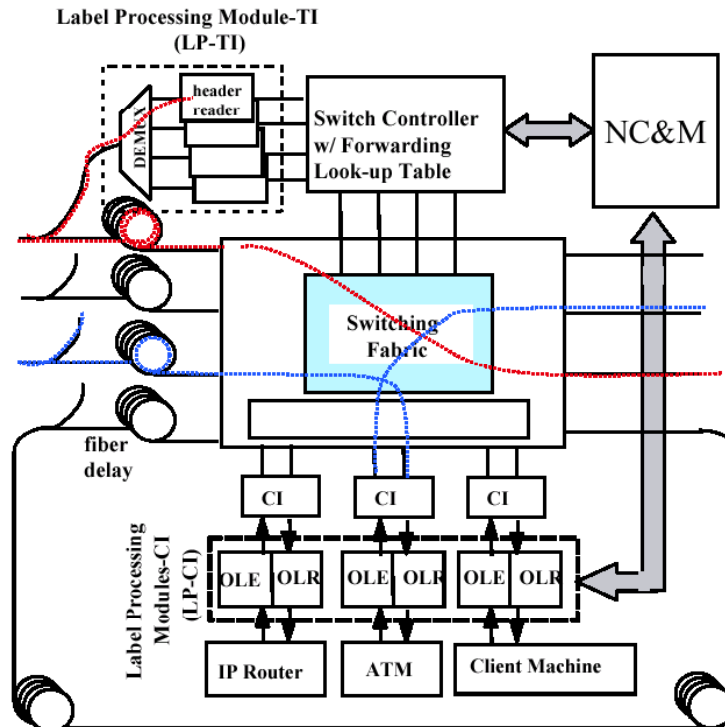
Figure 1: An example of an Optical Router with Label-Processing modules to achieve optical label switching for IP, ATM or client-specific traffic. Label-Processing Module at Transport Interface detects the signaling header for routing, and Label-Processing Module at the Client Layer includes Optical Label Encoder (OLE) and Optical Label Remover (OLR) to encode and remove the optical labels.

These headers or labels can be embedded within the packet or can be optically encoded on a separate optical subcarrier or another optical wavelength. While the proposed optical router can be configured to support any of the above protocol, the preferred method is optical-label switching using the optical subcarrier technique. New signaling information is added in the form of an optical signaling header which is carried within each wavelength in the multi-wavelength transport environment. The optical signaling header containing routing and control information such as the source, the destination, the priority, and the length of the packet, will propagate through the network along with the data payload. Each Optical Router will sense this optical signaling header, look-up the forwarding table, and take necessary steps to route the packet. During this processing and switching time, the packet is delayed by the optical fiber loop at the transport-input interface before entering the switch fabric. The goal is to reduce the need to manage the delay between the optical signaling header and the data payload if the network element itself provides the optical delay necessary for the short time required for setting the switch states within each network element. If the packet is to be routed to a wavelength/path where there is already another packet being routed, the Optical Router will seek routing by an alternate wavelength, by buffering, or by an alternate path. The optical subcarrier based optical-label switching has advantages over other methods in that the label and the data payload are attached to each other hence reducing the need to manage the time delay, and that the subcarrier header can be easily extracted without affecting the data payload.

In case of switching conflict, the Optical Router can take one, or a combination of the following three choices: wavelength conversion, time-buffering, and path deflection. Each has advantages and disadvantages, and we will compare and design optimum contention resolution algorithms.

- *Limited wavelength interchange,* where a packet may be converted to an idle wavelength that would allow it to be forwarded onto the next hop. We do not envision providing choices of an entire wavelength set as an alternative, however, providing a limited number of, at least one, alternative wavelength can effectively resolve contention. The alternative wavelength routing also **achieves the same propagation delay and number of hops** as the optimal case, and eliminates the difficulties in sequencing multiple packets. From this perspective, wavelength conversion is a very attractive solution compared to path deflection or time-buffering.
- *Limited delay buffering,* where a packet may be routed through a fiber delay line and recirculated back into an input port of the same Optical Router. At that point, the header content will be read and routing will be attempted again. One interesting point here is that, unlike in the conventional packet switching, no precise bit or packet synchronization is necessary for the buffer. For that reason, the choice of the length of the delay line can be arbitrary, however, there will be a trade-off between the contention resolution efficiency vs. the minimum optical latency. Limited number of delay buffer lines will be incorporated in the Optical Router, and multiple wavelengths are accommodated in each delay buffer line. Looping within the switching fabric is avoided by eliminating the possibility of a loop-back within the switch fabric.
- *Limited deflection routing,* where a packet may be deflected to a neighboring switching node from which it can be forwarded towards its destination. Care again must be taken to prevent a packet from being repeatedly deflected, thereby causing signal degradation, as well as wasting network bandwidth. The TTL and Loop prevention schemes discussed above can be applied here.
- *Prioritized packet preemption,* where a newly arrived packet may preempt a currently transmitting packet if the arriving packet has a higher priority. This would be the last resort solution when alternate wavelength, buffering, or path deflection routings are not available. Higher layer protocol such as TCP, or NC&M can request retransmission of the packet if necessary.

The generalized switching fabric architecture offers a number of key advantages over other architectures.

- It offers *space, time, and wavelength* switching by incorporating a tunable wavelength converter stage as the only reconfigurable stage, therefore, simplifying the control signaling architecture.
- It offers scalability and modularity far beyond the technological limitation imposed by AWGs
- For the same connectivity, the required AWG and tunable laser resolution is relaxed by the degeneracy factor *D*.
- Such relaxed resolution reduces the size of AWG by a factor of *D*, or improves the yield by a factor of D^2.
- The architecture allows modular scaling by adding a layer of AWG router at a time.
- The architecture supports wavelength conversion based contention resolution of degree *W/D*.
- Even with a current state-of-the-art AWG technology supporting 128x128, it is ultimately capable of achieving 42 Pb/sec switching capacity and 32x(65536x65536) connectivity.

Optical Wavelength Router
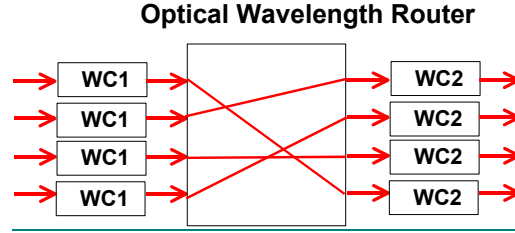
Figure 2:   The architecture of the core of the proposed optical router.

The architecture exploits the optical wavelength router which dissipates no power, and switching is achieved by tunable wavelength conversion at *WC1*. The power requirement scale as *2a'N* where *N* is the product of the total number of ports and wavelengths, and *a'* is a power rating of a wavelength converter. Each wavelength converter integrated with the diode laser will dissipate approximately 0.3 Watts (300 mA * 1 V) and can handle bit-rates exceeding 10 Gb/s. The total power requirements for a terabit optical router are projected to be below 60 W, and a petabit router to be below 60 kW. The switching fabric of Figure 2 is utilized in the optical router of Figure 1 so that this wavelength switching in WC1 accomplishes choice of output port (space), output wavelength (wavelength), and buffer line (time). Wavelength conversion is a powerful tool in optical routing since it accomplishes contention resolution without incurring additional latency or packet sequence skewing problems.

Figure 3 shows a schematic of the optical router switch fabric consisting of rapidly tunable wavelength converters (*WC1*), a wavelength router, and fixed wavelength converters (*WC2*). The first set of wavelength converters depicted as WC1 which will choose its wavelength that will route the signal to the appropriate output port of the wavelength router. The second stage wavelength converter WC2 will convert this back to the desired wavelength for the network. The wavelength router is based on the arrayed waveguide grating (AWG) which has the following well-known wavelength dependent routing characteristics.
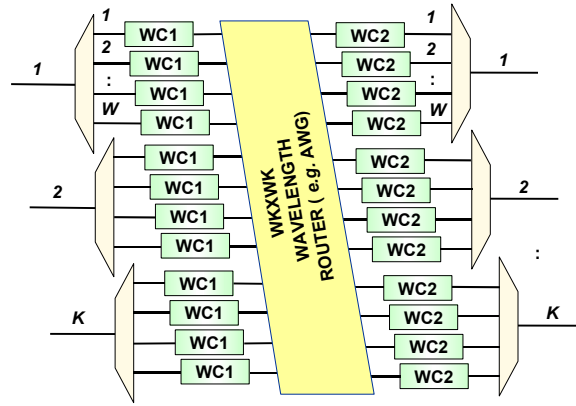
Figure 3:   A schematic of an optical router switching fabric for the NGI Supernet.

The Optical Router schematic of the original proposal is slightly modified as Figure 4. In this schematic, the optical label and the data payload are separated by the Fiber Bragg grating (FBG) and the optical circulator. The FBG achieves high reflectivities at the compliant wavelengths. The transmission through this FBG provides multiple optical-labels on multiple wavelengths which are subsequently demultiplexed and detected by simple photodiodes to recover header information on each wavelength. The reflection from the FBG provides the data payload at the compliant wavelengths which will be circulated and delayed by the optical fiber to provide time to process the header, to make a forwarding decision, and to tune the wavelength converter output wavelength. We envision this time delay would be less than 100 nsec, which can be brought shorter by advanced Application Specific Integrated Circuit (ASIC) designs. The switching fabric will be routed based on the decision made by the forwarding table, and will be routed to the output wavelength of the output port according to the decision. At that point, the packet will obtain a new optical-label written by an output wavelength converter WC2. The packet to be buffered will loop around a fiber delay line and will go through another decision and forwarding cycle. The forwarding decision will coordinate so that packets do not loop too many times. The packet can also be dropped to the client through a NCI-Out (Non-Complient Interface-Out). LP-CI-output consists of NCI-Out, Congestion Manager Out (CMo), and Optical Label Remover (OLR). Likewise, LP-CI-input consists of NCI-in, Congestion Manager In (CMi), and Optical Label Encoder (OLE). CMo will take appropriate actions to interface the packet to the LAN. CMi allows packets to be transported into the Supernet and coordinates with the NC&M to acquire an appropriate optical-label. Both CMi and CMo have electronic buffer memories and congestion management interfacing with the LAN and the Supernet.
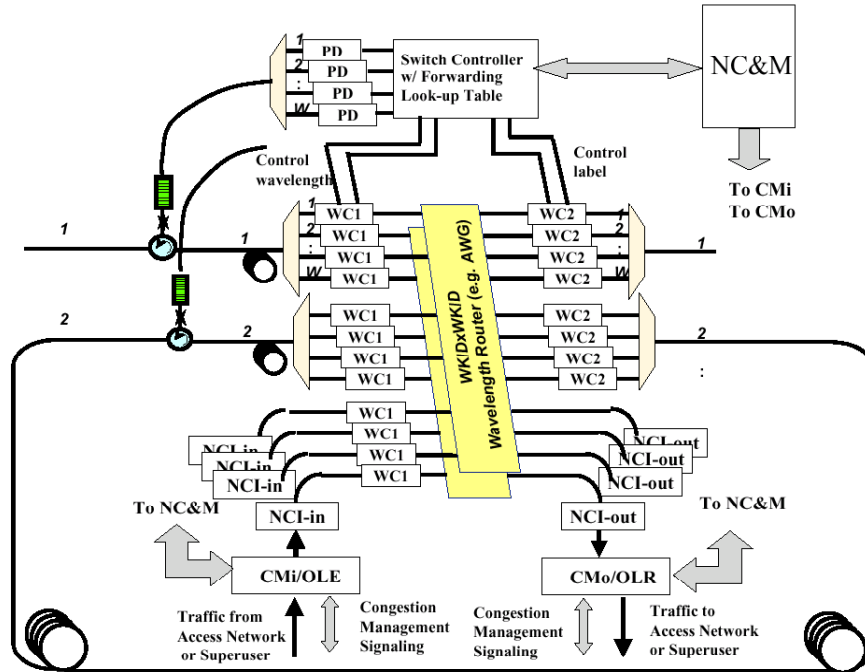
Figure 4: A schematic of the proposed Optical Router drawn in the context of optical-label switching network

4.2 Deliverable 1c: Optical Switching and Contention Resolution tradeoff evaluation complete (Y1Q3)

We have proposed, designed, implemented, and tested a novel contention resolution scheme which combines time, space, and wavelength domain switching. Further, this contention resolution scheme can apply to circuit-switching as well as packet-switching. The combined preference of contention resolution is as follows:

(1) wavelength conversion
(2) time buffering
(3) space deflection
(4) priority based pre-emption
(5) if none of the above solutions are available, drop the packet

Wavelength conversion is the most preferred packet contention resolution method because it does not induce any additional latency nor cause a skew in packet sequence. The second preferred method was time buffering. This is preferred over space deflection because of space deflection occupies the capacity of other parts of the network. When wavelength, time, and space based contention resolution is unavailable, there is no choice but to drop other packets of lower priority or to drop itself. Here, we use priority based pre-emption.

Network simulation using the above algorithms achieved highly efficient packet routing as shown in Figure 5. The proposed method applied to unslotted network can achieve similar performance as slotted network with fewer buffer numbers. However slotted network must segment natural IP packets and use synchronizers. The packet loss rate showed higher loss for lower priority packets due to differentiated pre-emption.
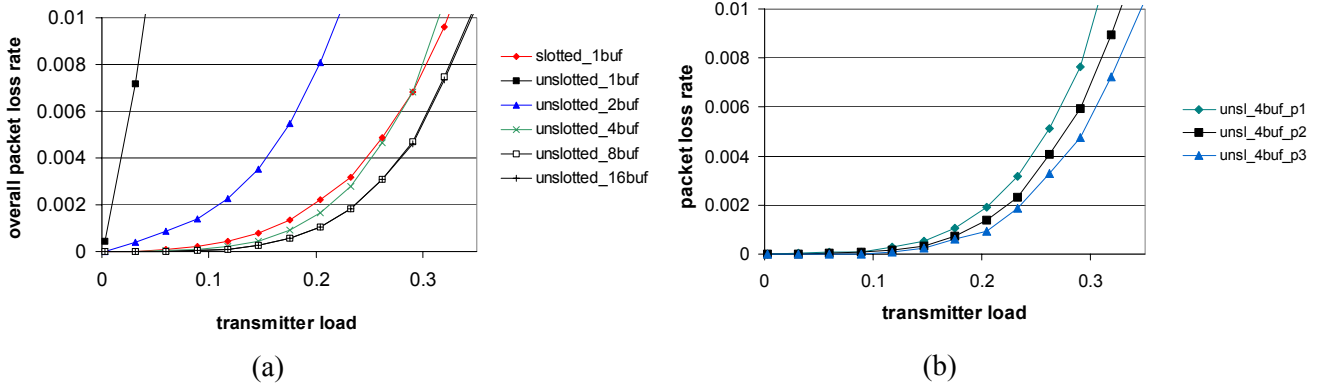
Figure 5: Simulated packet loss rate performace where (a) we compared performance of slotted vs. unslotted at various number of buffers, and (b) priority based preemption.

The network performance analysis is more accurate if self-similar traffic is used to imitate bursty traffic. The modeling of self-similar traffic is based on superposed on-off sources (Figure 6). The *on* and *off* periods of each on-off source are modulated independently by Pareto distribution. Pareto distribution has the following probability density function:

$$P(x) = \frac{\alpha b^{\alpha}}{x^{\alpha+1}} \quad , \quad x \geq b \tag{1}$$

Where α is a shape parameter (tail index) and $1<\alpha<2$, and b is minimum value of x. The Hurst parameter of the traffic is given by:

$$H = \frac{3-\alpha}{2} \tag{2}$$

During the *on*-period packets arrive back-to-back. The packet size distribution is modeled after the real IP data collected by CAIDA (www.caida.com). Figure 7 shows the cumulative probability distribution of IP packets. In the simulation a 19-order polynomial was used to produce a closely matching distribution. The maximum packet size is set at 12,000 bits.
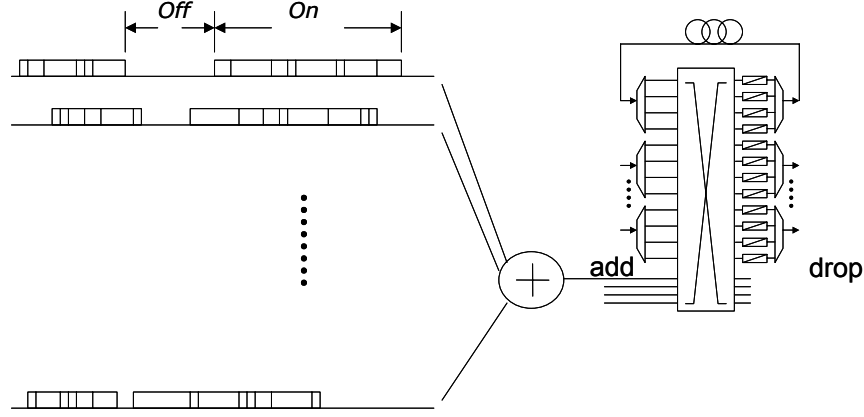
Figure 6: On-off source based Pareto distribution packets.

packet size (bytes)
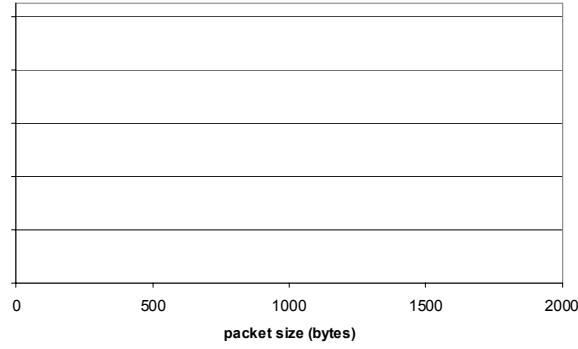
Figure 7:   Cumulative probability of packet distribution measured over CAIDA

In the simulation, 12 on-off sources contribute to each transmitter. Each on-off source has OC-3 data rate, while the transmitter has OC-48 data rate. The *on* period is modulated by Pareto distribution with α=1.4, the *off* period is modulated by Pareto distribution with α=1.2.

Electrical buffers are used at the ingress nodes. That is, before injecting the packet into the optical switch, the control unit will first inspect if the wavelength (which the packet is using) at the output port is busy or not. If that wavelength is busy, the packet will be held until that particular output port clears. Although it creates head-of-the-queue blocking, the results show that it can significantly reduce packet loss rate.

Figure 8 and Figure 9 show the packet loss rate and end-to-end delay of different configurations. There are four wavelengths in each fiber. At each switch the wavelength conversion degree is set at 2. During contentions, deflection is chosen before using preemption. By using electrical buffers at the ingress nodes, we were able to achieve packet loss rate of less than 1% at 40% transmitter load. This approach does not introduce significant queuing delay because the transmission of one packet is very small, compared with the propagation delay.
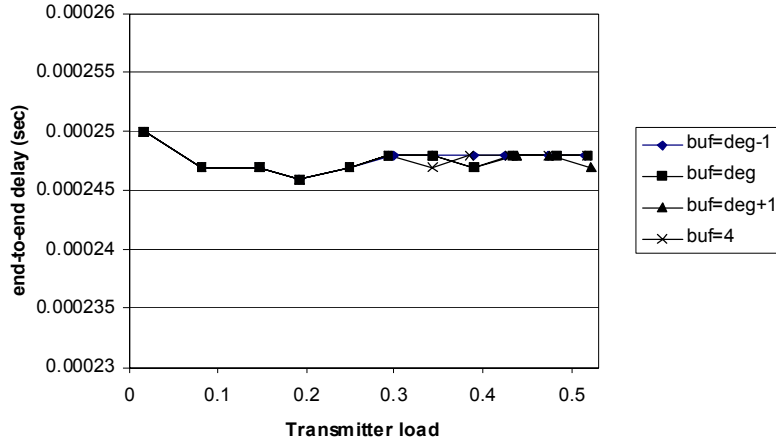
Figure 8: Packet Loss Rate for a varying number of buffer numbers.
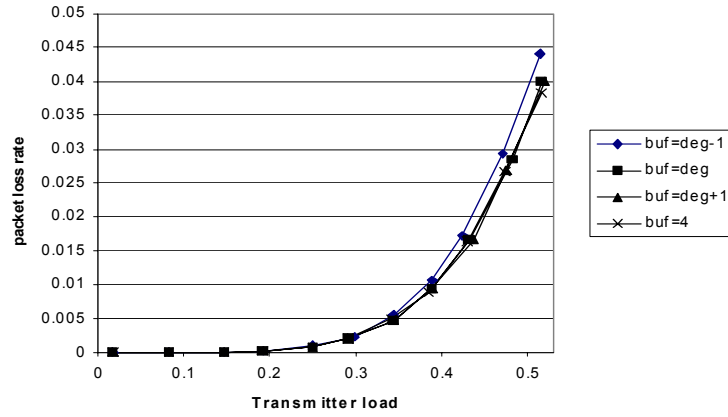
Figure 9: End-to-end delay for a varying number of buffer numbers.

4.3 Deliverable 1b: Router Interoperability Study of Optical-Label Switching, Optical-MPLS, Terabit-Burst Switching, and Optical-Flow switching (Y1Q4)

One of the key ideas behind the Optical Router is to introduce wavelength, time, and space domain contention resolution rather than the conventional method relying primarily on the time domain buffering and queuing. This new architecture provides a great degree of flexibility for the optical router to work with a broad range of networking protocols, and offers very little latency and the timing jitters. The introduction of two-tier signaling, in-band (subcarrier multiplexing) and out-of-band (data communication channel at 1510 nm) signaling, further facilitates the interoperability. Traditional WDM networks utilize network control and management signaling at 1510 nm Data Communication Channel. The optical label, implemented in the subcarrier of the data payload wavelength, provides signaling useful for packet switching. Using the two-tier signaling, the Optical Router can achieve Optical-Label Switching, circuit-switching, burst-switching, flow-switching, and packet switching. Figure 10 shows the two-tier signaling and the logical processes leading to forwarding decision and routing functions.
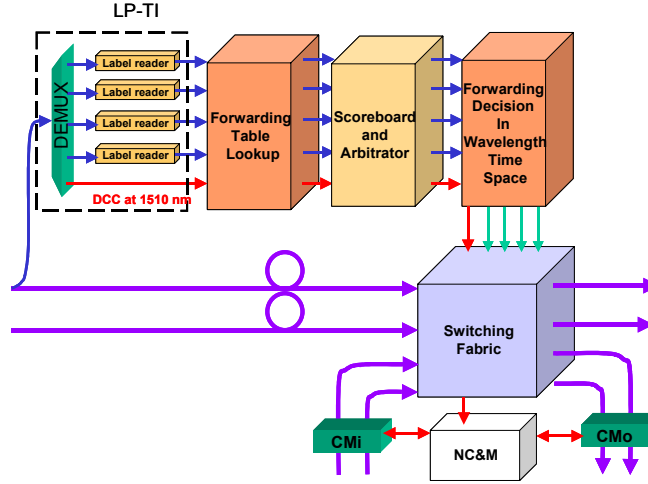
Figure 10:   Signaling for forwarding decision functional block

A circuit-switching implementation of the Optical Router will utilize signaling via the data communications channel at 1510 nm for provisioning the circuit. The NC&M system implemented on this DCC will provide signaling between the optical routers, as well as User to Network Interface. It is important to note that the signaling via DCC should be interfaced to the forwarding table and the score board (the unit which keeps track of which channel of which port are occupied), so that the contention resolution can be achieved including the established circuits.

A packet-switching implementation of the Optical Router will require signaling via the subcarrier optical-labels, in which in-band signaling information related to packet forwarding can be quickly read, interpreted, compared against the forwarding table, arbitrated according to the scoreboard, and eventually induce the forwarding decision. This time plus the switching time of the switching fabric will be implemented in the form of the fiber delay line. Due to its asynchronous and unslotted architecture, the Optical Router allows interoperability with burst switching and circuit switching without the need for segmentation, synchronization, and reassembly. Moreover, the Optical Router allows greatly simplified switching fabric without the synchronizer.   Introduction of the new wavelength domain contention resolution by wavelength conversion allows lower latency, lower timing jitter, and reduced need for packet resequencing.

A burst-switching implementation of the Optical Router is also relatively simple. Since optical burst switching can be seen from two different perspectives, the implementations can also be in two different configurations. The first perspective is to view an optical burst as a short duration of a circuit, and optical burst switching requires unidirectional reservation of resources for a short duration. The second perspective is to view an optical burst as a long packet, and optical burst switching requires signaling that is no different from that of optical packet switching described above. For short bursts (sub-milliseconds), these bursts can be considered as relatively large packets. The burst control packets are the labels carried on subcarrier on each wavelength along with the data payload. The label structure will be identical to the packet switching case except that the packet duration is programmed to be the burst duration.

A MPLambdaS implementation of the Optical Router is achieved by accommodating either CR-LDP or RSVP signaling on the DCC, and switching the Optical Router switching fabric according to the signaling in order to provision a light-path circuit. Note that, in this case, the Optical Router works exactly like an Optical Crossconnect.

A flow-switching implementation of the Optical Router is best implemented via the RSVP signaling discussed earlier. As in the case of MPLambdaS, the RSVP signaling on DCC informs the Optical Routers about the flow that is about to initiate. An alternative and more dynamic reservation is achieved if the table look up and the scoreboard referencing steps themselves can trigger initiation of the flow switching based on similar pattern of connections that have been repeatedly requested.

Table 1 compares the IP-rsvp, the Optical Burst Switching and the Optical-Label Switching technologies from several perspectives. An important point worth reiterating here is that the Optical Router can accommodate mixed protocols of the above on multiple wavelengths and ports. This can be especially useful in today and future networking where diverse needs of networking (circuit and packet ; high priority and low priority; time-sensitive and time-tolerant applications) are present.

TABLE 1: Comparison of IP-RSVP, Optical Burst Switching and Optical-Label Switching

Protocol	Conventionial IP-rsvp	IP-rsvp over MPLS	IP-rsvp over MPλS	OBS	OLS
Signaling	IP rsvp used to reserve resources (buffer, etc)	IP rsvp used for signaling and LSP setup	IP rsvp used for signaling and lightpath setup	Burst control packets used for burst setup	Both DCC (out-of-band) and subcarrier label (in-band) used for signaling
Reservation direction	Both Unidirectional reservation and Bi-directional reservation possible	Bi-directional reservation with LDP and LSP setup	Bi-directional reservation with LDP and lightpath setup	Unidirectional reservation only	Both bi-directional reservation (e.g. circuit) and Uni-directional reservation (e.g. packet) possible
QoS parameter	QoS information carried on the 'PATH' message	QoS information mapped implicitly to the Label	QoS information mapped to the Label (here QoS is more related to the quality of lightpath, e.g. reliability, restoration time, etc.)	QoS information controlled by the time offset between the burst control packet and the burst	QoS information carried on DCC for guaranteed QoS, or on the subcarrier label for priority based CoS
Design Purpose	Designed for switching micro-flows	Designed for switching virtual circuits or flows	Designed for switching hard circuits (wavelengths)	Designed for switching bursts (granularity is not yet defined)	Designed for switching circuits, bursts, flows, or packets
Required Hardware	Assumes an electronic router with buffers and RSVP enabled	Assumes an electronic MPLS router with buffers	Assumes an electronic router working with an optical crossconnect	Assumes an optical crossconnect with a burst processor.	Assumes a fast optical crossconnect with a label processor

4.4 Deliverable 2a: Optical-Label Header Processors Operational (Y2Q1)
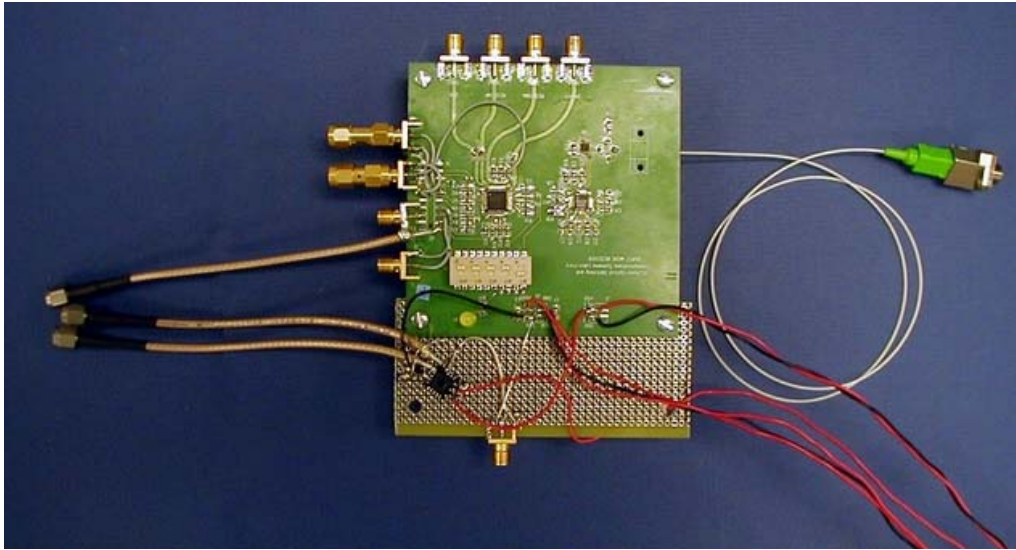
Figure 11: Burst Mode Receiver

Figure 11 shows a picture of designed burst mode receiver to process incoming optical label signal. We have successfully demonstrated all optical packet switching experiment using the optical label header processor. And also the burst mode receivers have been functionally field-tested over Sprint/NTON fiber infrastructure using packetized header signal from the same source. During the field-test experiment, the FIFO buffer depth was set to 7 bits (instead of 5 bits that was used in the lab environment) to eliminate the jitter associated with the input signal from the FIFO output clock and data signals along the field-tested fiber. This is illustrated from the functional diagram of the burst mode receiver, which is shown in Figure 12. The buffer adjustment introduced an additional 1 to 2 bits delay to the synchronization process, which translated into a 10 to 20 ns delay for the header recovery process that operates at 100Mbit/s (NOTE: The data delay is 3 to 4 bits for 7 bit FIFO buffer depth and 2 to 3 bits for 5 bits FIFO buffer depth).

For a more thorough performance characterization and improvement, the following tasks should be performed.

- **Guard time:** Guard time is defined as the time between two arrival packets. By varying the guard time between packets, it allows testing the receiver's threshold-adjusting circuit and stresses the clock-recovery circuit by applying odd guard times.

- **Bit rates:** Due to the processing speed limitation on the FPGA, the burst mode receiver is currently operates at 100Mbit/s, which is the minimum bit rate that the receiver can operate. For typical operation, the bit rate should be increased to 155.52 Mbit/s.

- **Reset Circuitry**- A reset pulse is need for (1) resetting the threshold of the Clock and Data Recovery (CDR) chip and (2) clock synchronization. Without the reset pulse, the automatic threshold control within the burst mode receiver may fail to detect a low optical power level packet that closely follows a high optical power level packet. Furthermore, a much higher data transition density is needed to prevent DC walk-off as there could be a long idle time between arrival packets. With the reset pulse, the data transition density can be as low as 1/72 [transition/bits]. Besides providing reset to the threshold level, a reset pulse is also required for proper clock synchronization between the local clock and the recovery clock. According to the spec sheet, the reset pulse should be asserted 2 bits or more prior to data arrival, which is shown in Figure 13.
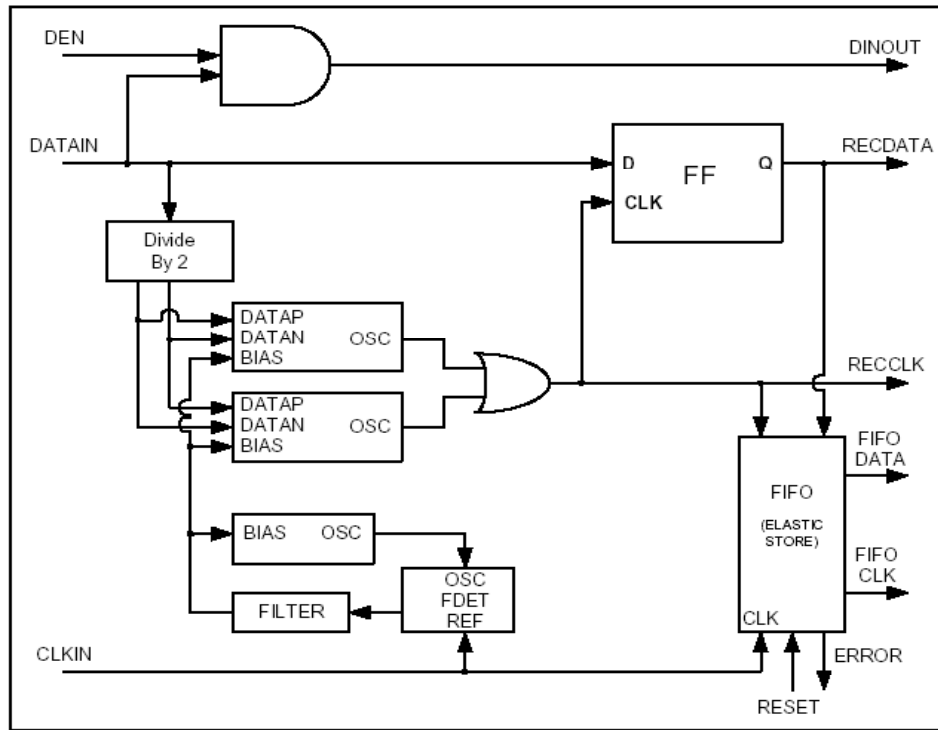
Figure 12: Functional Diagram of Clock and Data Recovery Chip

Currently, the burst mode receiver is sufficient to operate at header bit rate of up to 180Mb/s. However, as the header bit rate scale up to OC-12 (622.08 Mbps) or even higher bit rate to accommodate the growing demand of switching functionalities (i.e. more header information is achieved by operating the header at a higher bit rate with the same time duration), alternate technologies, such as Gigabit Ethernet, maybe needed. The benefit of using Gigabit Ethernet is that it is compatible with the existing Ethernet. In term of Open Systems Interconnection (OSI) layering infrastructure, Gigabit Ethernet will look identical to Ethernet from the data link layer upward. This provides a graceful transition from the experimental protocol to a standardized protocol. However, there are many issues that should be investigated:
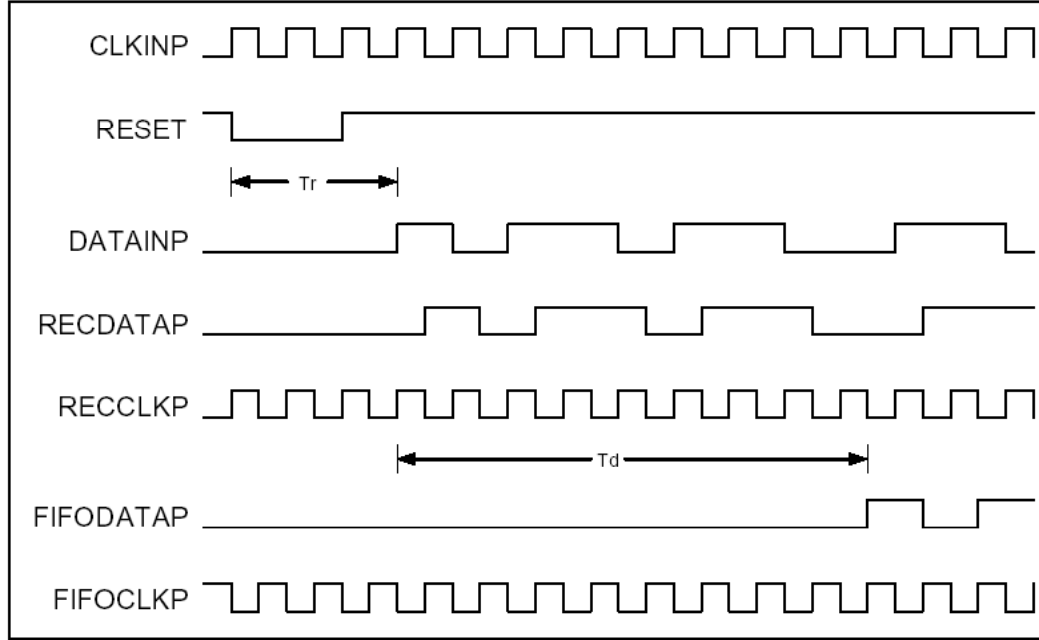
Figure 13: Timing Diagram of Clock and Data Recovery Chip

Frame Structure: The Ethernet frame structure is shown in Figure 14. If the header information were to be encapsulated in the Data field, the minimum size of the header information will be restricted to 46 bytes, or 368 bits. Although the data field can be "stuffed" with unused bits to fill it out to 46 bytes, it might not be practical for our application. Since the Destination Address field and Source Address field are not needed for our application, the header information may possibly be encoded in the destination or the source address field. Further studies are needed to find an intelligence way to utilize the frame structure.

Modulation: Since Ethernet employs Manchester encoding, the bandwidth needed for the header signal will be two times as much as compared to On-Off Keying (OOK) that we used previously. The Sub-Carrier (SCM) transceiver might need some optimizations such as subcarrier frequency adjustment to accommodate the change.

Preamble (8 bytes)	Dest. Addr. (6 bytes)	Source Addr. (6 bytes)	Type (2 bytes)	Data (46 – 1500 bytes)	CRC (4 bytes)

Figure 14: Ethernet Frame Structure

4.5 Deliverable 3a: Header Processor with Forwarding Table Integration Complete (Y2Q2)

This summary describes the progress made on the design and implementation of a header processor with forwarding table integration for an optical label switching (OLS) router. The aim was to build a header processor with forwarding table lookup that supported OC-3 label rates with number of fibers, $N = 4$ and number of wavelengths/fiber, $W = 4$, with minimum forwarding latency, while the payload bit rate could be anywhere from OC-12 to OC-192 rates. The advantage of using sub-carrier multiplexed (SCM) label

and payload is that as long as there is no cross talk between their frequency spectrums, their bit rates are independent of each other. Since the header processor is only concerned with the label bit-rate, this was a big design advantage compared to both high-speed electronic and optical routers. Additionally, the asynchronous and unslotted packet forwarding requires no complex segmentation and reassembly of packets at high payload rates. A commercial off-the-shelf FPGA was used as the header processor with forwarding table lookup, to satisfy the packet forwarding requirements. Further, the FPGA based processor allowed experimenting with different forwarding strategies and arbitration techniques and thus considerably reducing the design and verification cycles. The flexibility of reconfigurable logic allowed us to adapt quickly to changes in receivers or tunable lasers or switch fabric and not be limited to specific vendors, while avoiding expensive ASIC redesigns. The header processor has been implemented using a single Xilinx VirtexE family (XCV1000E-BG560) of FPGA. Fig. II.D.5.1 shows the top view of the implemented header processor board.

Figure 15: Implemented header processor with forwarding table lookup board

The burden of forwarding packets and resolving output wavelength and port contention is placed on the header processor. The processor performs forwarding table look-up based upon the label contents and after fair arbitration of the output ports, sends control signals to subsequent tunable wavelength converters (T_WC). Based on the control signals, the T_WC tunes to the appropriate wavelength and forwards packets through the switch fabric.

Figure 16 shows the block diagram of the header processor and indicates the three main components, *IChannel*, *arbiter* and *scoreboard*, in the architecture. At the router, the label is extracted from the payload and each extracted label is converted into electronics by its corresponding burst-mode receiver (BM_RX) and fed to the processor aligned to a master clock.

The first stage in the processor architecture is called the *IChannel*. Each label input from the various BM_RXs has a unique *IChannel* assigned to it, which operates independent of other *IChannels*. The reason for the independent operation between various *IChannel*s is that packets arrive asynchronously at the router. Hence, the corresponding labels have to be processed independent of each other's arrival instant.
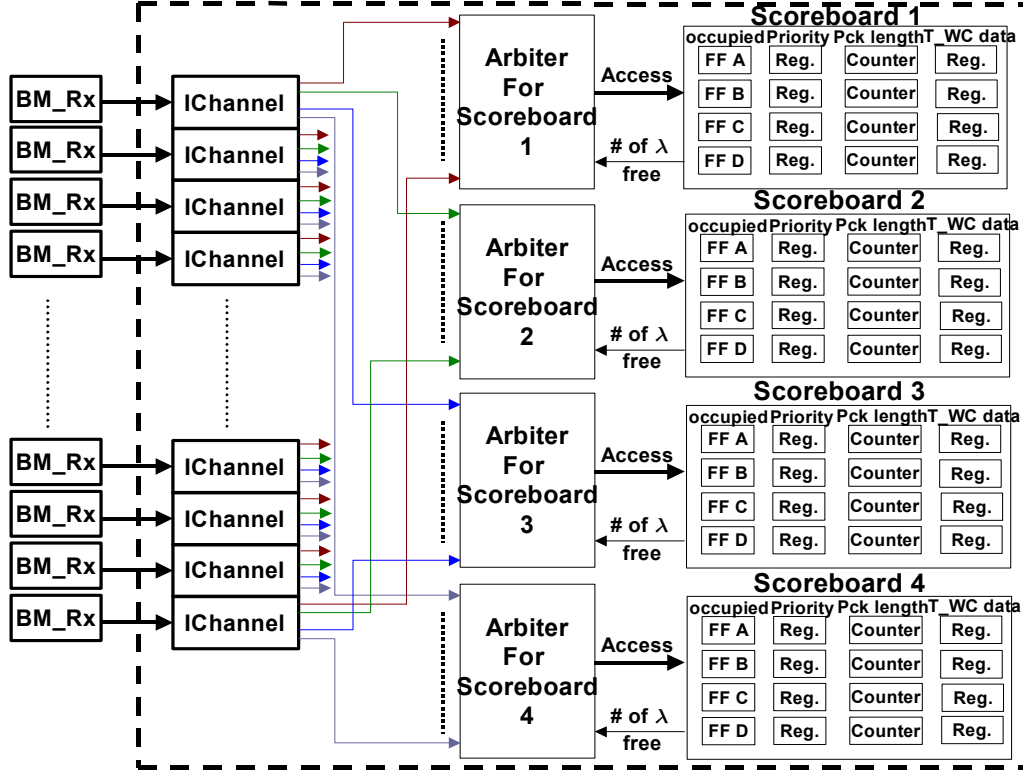
Figure 16: OLS router header processor internal architecture

The most important section in the processor architecture is a global structure called the *scoreboard*. The *scoreboard* consists of state-machines, counters, and registers that together inform the *arbiter* whether a wavelength on a particular output fiber is free or occupied. Thus, each *scoreboard* represents and monitors all the wavelengths on a unique output fiber. The access to each *scoreboard* is regulated by means of the *arbiter* that accepts requests to be forwarded through the corresponding *scoreboard* if and only if the *scoreboard* indicates that there is a free wavelength on the corresponding output port.

The current header processor implementation supports 4 optical fibers with 4 wavelengths on each fiber. This maps to 16 *IChannels* and as a result 16 requestors on each of the 4 *arbiters*. The critical path in the design is through the *arbiter*. This 16 input *arbiter* has a register-to-register delay of 11.92nsec (36.3% logic delay and 63.7% routing delay) with 10 levels of fanout-of-four (FO4) gates. The current implementation supports OC-3 label bit-rate input and operates internally at 78.8MHz. If this design is implemented as an ASIC in 0.13μm technology, it is reasonable to expect that the header processor can operate at around 650MHz, with the resultant 9-fold improvement in the latency measured in absolute time.

4.6 Deliverable 2b: Arrayed Waveguide Grating Operational (Y2Q3)

- *Fabrication of arrayed waveguide gratings (AWG)*

In the fabrication of the device, a lot of time was spent in determining the ideal chemistry to obtain smoothened sidewalls that have a good performance in terms of waveguiding. A variety of InGaAsP/InP etching steps were tried out.

1. Wet etching (in different proportions of Saturated Bromine Water (SBW): HF: DI Water). This was done in order to determine the ideal proportion in which the etching did not last too long, as this would smoothen out the sidewalls too much. A 1:2:2 gave a good profile for the sidewalls although it did not contribute to reducing the etching time substantially.

2. Dry etching (Chlorine and CH_4/H_2) – Chlorine based etching was carried out in Korea. The SEM pictures are attached. This had a detrimental effect on the sample in its roughness and had to be wet etched for a few minutes to smoothen out the walls. Samples of AWGs have also been sent to the Royal Institute of Technology in Sweden for dry etching using CH_4/H_2. This gives a better sidewall characteristic and a subsequent wet etching would improve the overall quality of performance of the device.

- *OMCVD Regrowth*

The regrowth was carried out at the Lawrence Livermore laboratory (LLNL) after the wet etching step and also at Sweden where dry etching is carried out first followed by the regrowth step. Results are awaited from Sweden. The SEM pictures from the LLNL regrowth have been attached Figure 17 below.
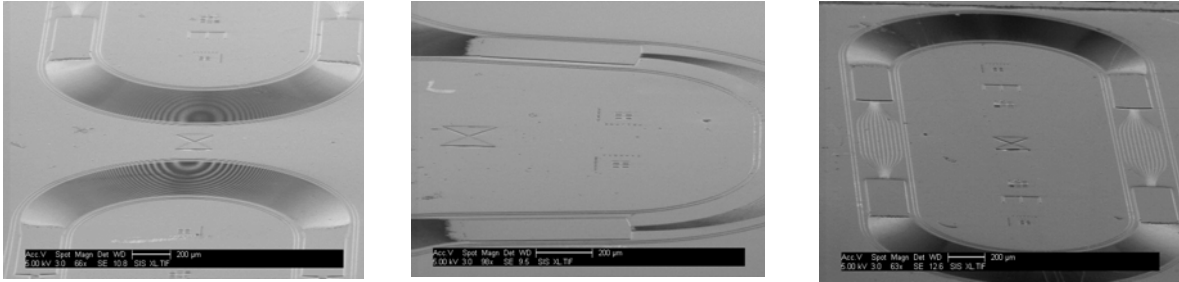
Figure 17:   An SEM picture
showing Arrayed Waveguide Grating sample after the regrowth step.

- *Device Testing*

The characterization of the AWG is being carried out at the moment. The samples obtained after wet etching and the samples after wet etching followed by regrowth are being tested. The testing set up consists of a stage (that can be adjusted in its movements in the x, y and z directions) on which a lens fiber couples light into the waveguide and the light is coupled out with another lens fiber. Lens fibers are used as they aid in matching the mode of the light launched in the fiber to the light coupled in the waveguide and vice-versa. The testing is carried out for straight waveguides (of different lengths), curved waveguides and finally the arrayed waveguide grating itself. The testing results are still being compiled and should be ready in a short time. This enables one go back re-check what corrections need to be made in the design and fabrication in order to obtain a better performance the next time. The current design is based on InP/InGaAsP waveguide material system to facilitate future 3-D integration Effective index method based waveguide design is performed.

The second generation AWG router must incorporate loss compensation over wide spectral range regardless of the choice of N x N paths. This requires a far more sophisticated design than that of the first generation device, and we are collaborating with Dr. Katsunari Okamoto at NTT. Currently, NTT group is the only group that has demonstrated 'Uniform Loss Cyclic Frequency (ULCF)' AWG routers. Through collaboration, we are currently designing UCLF AWG routers to be integrated in three-dimensions. Figure 18   (a) indicates the unique mapping plane with 3 dB coupling that allows compensation of port dependent loss. The AWG has eight input and eight output ports, and incorporates a special design to equalize optical loss for all 64 paths and to operate cyclically over the wide optical frequencies. Figure 18   (b)   shows measured AWG spectrum using the 8x8 ULCF AWGs obtained from

Dr. Okamoto via a collaboration arrangement. The first input port is used and respective output spectrum at each one of the eight output ports were measured and overlaid. The measured peak transmission was **uniform at -8.3±0.5 dB** over the entire spectrum. The 3 dB coupler internal to the AWG router gives intrinsic loss of 3 dB.

(a)   (b)
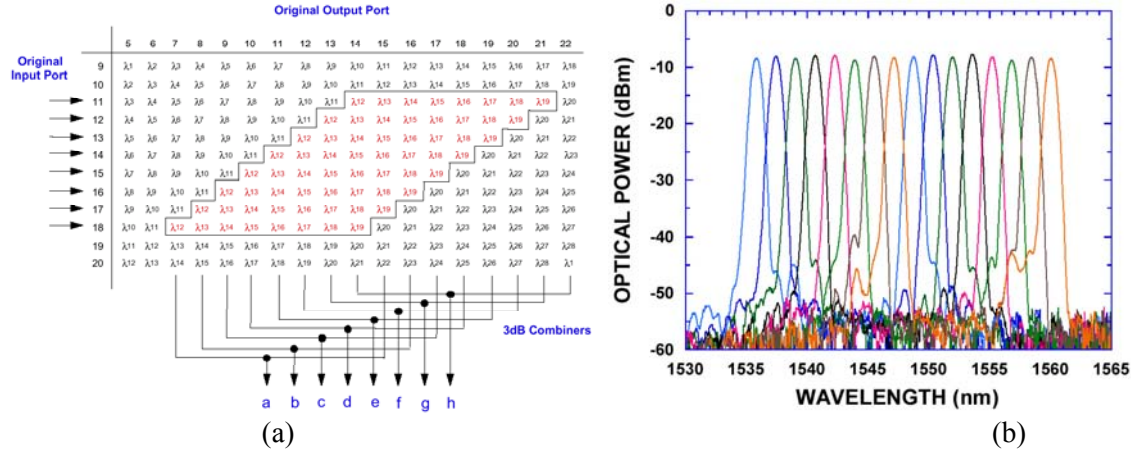
Figure 18: (a) Unique mapping plane that allows Uniform Loss Cyclic Frequency (ULCF) AWG characteristics. Note the 3 dB couplers in the AWG that provides loss compensation. (b) measured transmission spectrum of the actual ULCF AWG.

4.7 Deliverable 2c: Rapidly Tunable Wavelength Converter Operational (Y2Q4)
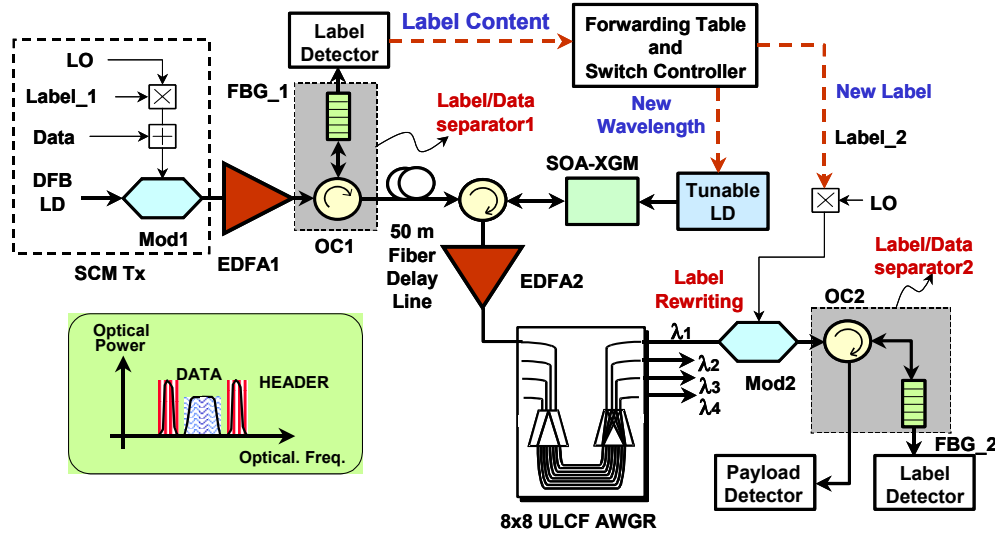
Figure 19: A functional block diagram of the experimental system, which represents the core of an Optical Label Switching (OLS) Router. The experimental setup includes optical circulators (OC1 and OC2), fiber Bragg gratings (FBG1 and FBG2), Erbium-doped fiber amplifier (EDFA1, EDFA2), LiNbO₃ modulators (Mod1, Mod2), 14 GHz local oscillators (LO), a polarization controller (PC), and photodetectors (Payload detector and label detector).

Figure 19 shows a functional block diagram of the experimental system, which represents the core of an Optical Label Switching (OLS) Router. It consists of an optical-subcarrier multiplexing transmitter, an optical-label/data separator, a label detector, a forwarding table, a switch controller, a tunable wavelength converter including a tunable laser and a semiconductor optical amplifier (SOA), a uniform-loss-cyclic

frequency (ULCF) AWGR, label rewriters and receivers. The transmitter modulates the original label (label_1) at 622 Mb/s on a 14 GHz carrier frequency subcarrier multiplexed with the baseband data payload packet at 2.5 Gb/s. The modulated signal will then include double-side-band subcarrier header appearing 14 GHz away from the center optical frequency of 192.7 THz (1549.3 nm). The optical-label/data separator consists of an optical circulator (OC1) and a Fiber Bragg Grating (FBG) with its peak reflectivity centered at the same optical frequency. The FBG has higher than 99.9 % peak reflectivity with a pass band below 5 GHz Half-Width-at-Half-Maximum. Hence the FBG nearly perfectly reflects the data payload appearing at the center lobe and almost completely transmits the subcarrier header appearing at the side bands. The combination of the optical circulator and the FBG separates the input signal into the data payload and the subcarrier header. The data payload will be over-modulated with the new label content at the output of the OLS router to achieve all-optical label-swapping. The header detector demodulates the 40 bit long optical-label, which will be compared against the content of the forwarding table, and, based on this comparison, the controller will generate a control signal for the switch fabric and a new optical-label content for the label rewriter (Mod2). The switch control signal will instruct the rapidly tunable laser to tune to the wavelength corresponding to the desired packet output destination port of the AWGR. This entire process from the label detection to the optical switching takes 250 nsec, which matches the optical delay time of the 50 meter optical fiber for the data payload. After this delay, the data payload enters the SOA and modulates its gain. This gain modulation (XGM) will then modulate the intensity of the counter-propagating emission from the tunable laser, which will now assume the inverted modulation information of the data payload. The counter-propagating geometry avoids the need for using an optical filter against the transmitting tunable laser output, and also provides additional bandwidth-limit to further reduce the leakage of the subcarrier optical-label signal at the higher modulation frequency. The inversion of the data payload modulation could be corrected by cascading another inverting wavelength converter in tandem at the output port of the OLS router.

The experiment utilized two optical-label contents for two output wavelengths λ_1 (1547.22 nm) and λ_2 (1554.22 nm). The tunable laser covers a 45 nm tuning range between 1525 and 1570 nm by adjusting the current applied to the three electrodes. The AWGR has eight input ports and eight output ports with 200 GHz channel spacing, and incorporates a special design to equalize optical losses for all 64 paths and to operate cyclically over the wide optical frequencies. This power equalization across all switching states in the fabric is extremely important for all-optical packet switching systems, especially when power equalization and optical regeneration are considered. At the output ports of the AWGR, additional optical modulators over-modulate the new subcarrier label on the data payload, and hence achieving optical-label swapping. Dynamic optical-label switching and packet routing experiments with optical label switching were conducted by using two types of optical labels, label_1 and label_2, each of which dictates forwarding the packet to output port_4 and port_8 of the AWGR by tuning the laser to λ_1 and to λ_2. The timing diagram is similar to that in Figure 21 (a) except that the data payloads attached to label_1 and label_2 are much shorter and asymmetric at 320 nsec and 930 nsec durations respectively in order to simulate asynchronous and unslotted packet routing. To further analyze the dynamic characteristics of the OLS routing system, each output of the OLS router was monitored by a baseband optical receiver with a bandwidth limit at 3 GHz (marked as payload detector in Fig. II.D.7.1).
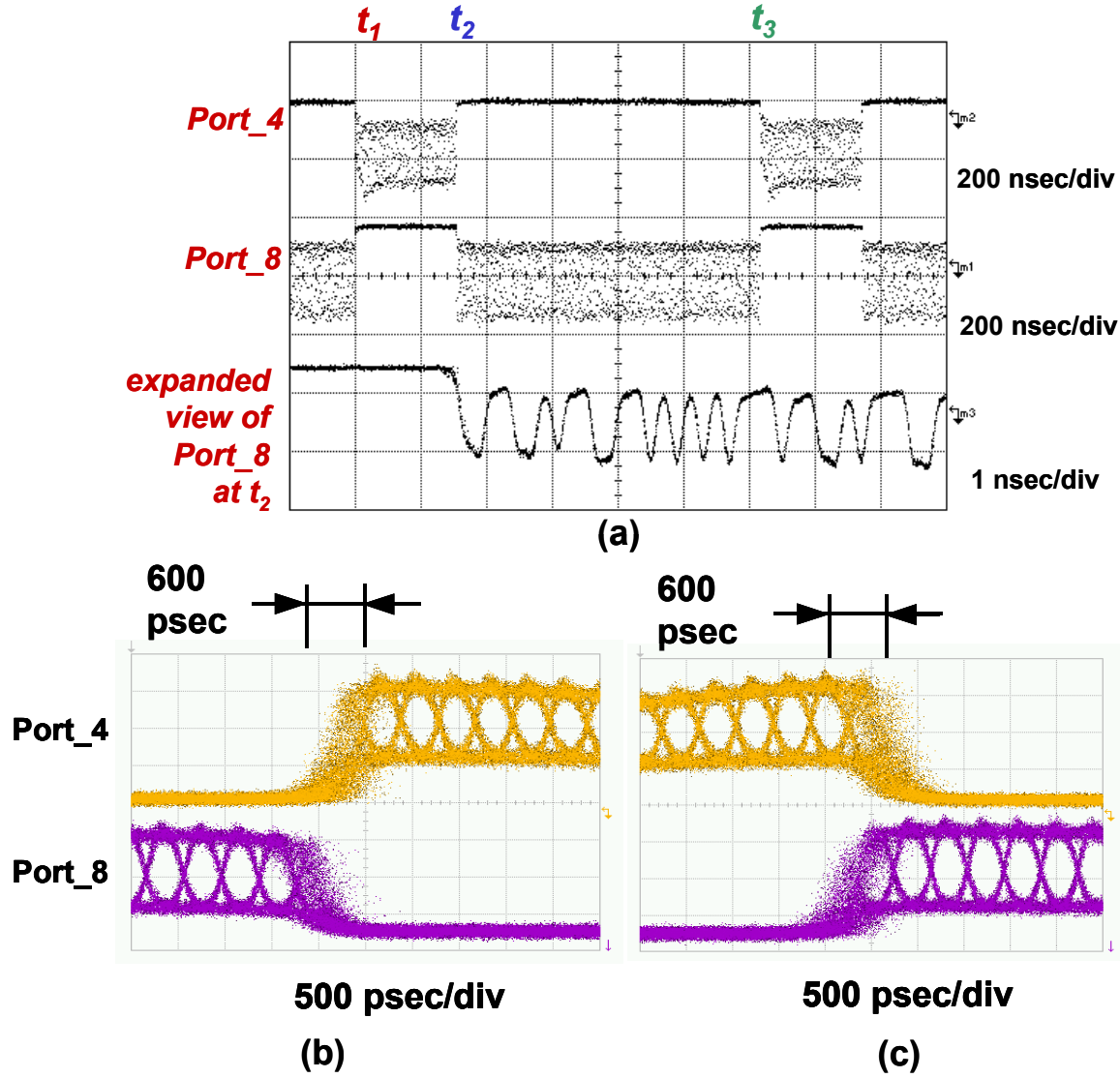
Figure 20:   (a) The optical signal detected at port_4 (top trace) showing packets routed between t_1 and t_2 , and the signal detected at port_8 (middle trace) showing packets routed between t_2 and t_3 on a 200 ns/div scale. The bottom trace shows the expanded view of the middle trace near t_2 on a 1 nsec/div scale indicating clear 2.5 Gb/s bit patterns with a negligible pattern dependency. (b) expanded views around the t_1 region on a 500 psec/div scale detected by inverting photoreceivers. (c) expanded views around the t_2 region on a 500 psec/div scale detected by inverting photoreceivers.

Figure 20   (a) displays the optical signal detected at port_4 (top trace) showing packets routed between t_1 (200 nsec point) and t_2 (520 nsec point), and the signal detected at port_8 (middle trace) showing packets routed between t_2 (520 nsec point) and t_3 (1440 nsec point).   The bottom trace shows the expanded view of the middle trace near t_2 indicating clear 2.5 Gb/s bit patterns with a negligible pattern dependency.   The top and middle traces are on a 200 nsec/div scale, and the bottom trace is on a 1 nsec/div scale. Fig. II.D.7.2 (b) and (c) show further expanded views on a 500 psec/div scale around the t_1 region and the t_2 region respectively detected by inverting photoreceivers and using a pseudorandom data bit sequence of 2^{15}-1 word length. The clear eyes and the very rapid transition times of 600 psec are evident. This transition time is limited by the combination of the driver circuitry and the tunable laser

dynamics. The above experiments utilized no guard time in the data payload in order to manifest the transitions in the packet switching. Packet bit-error-rate (BER) measurements incorporated 500 nsec guard time and 100 µsec long data packets under both label_1 and label_2, and accumulated packet BER measurements conducted at Port_4 and Port_8. Figure 21 (a) and (b) show the timing diagram and the accumulated packet BER measurement result. Both output ports showed error free operations. Slightly different power penalty levels for the two ports are mainly due to the wavelength dependent characteristics of the SOA wavelength converter operating differently at λ_1 and λ_2. The inset eye diagrams are for the optical-label (upper trace) and the data payload (lower trace) at the output of the OLS routing system after the label swapping. Both eye diagrams show clear openings.
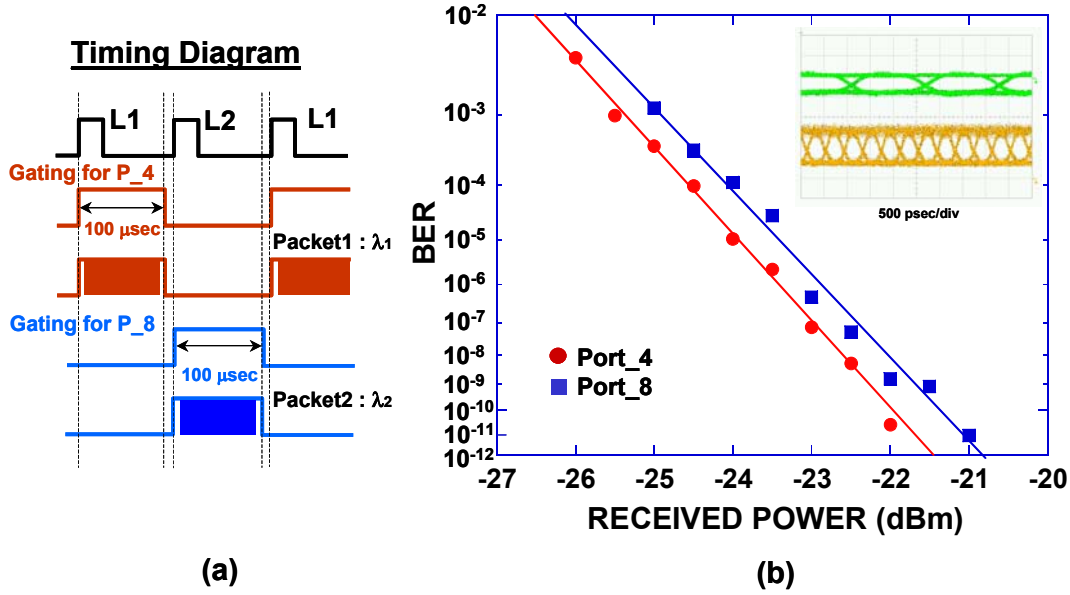
Figure 21:   (a) The timing diagram for the packet BER measurement. (b) the accumulated packet BER measurement result.

The demonstrated packet routing system is potentially capable of 10 Gb/s data rates limited by the 10 GHz full-width half maximum of the FBGs. Higher-bit rates can be accommodated by using higher optical intensity in the XGM SOA and single-side band optical-labels, or by using parametric wavelength conversion processes and time-division-multiplexing optical-labels. Further, incorporating wavelength conversion technologies with 2R or 3R regeneration capabilities such as Mach-Zehnder interferomic wavelength converters can greatly extend the cascaded operation limit of the optical router system. 10 Gb/s and higher data rate packet routing experiments with optical regeneration incorporating a 50 GHz channel spacing ULCF AWGRs are currently in progress.

4.8 Deliverable 3b: Controller and Driver Integration Complete (Y3Q1)

As shown in Figure 22, the function of the tunable laser and driver is to provide laser power at a desired wavelength for the wavelength converter according to the control signal from the FPGA which is the switching control. Combined with the wavelength converter they form a tunable wavelength converter subsystem capable of tuning fast for the packet switching application. The basic design of the driver is as following: The tunable laser consists of 4 sections – gain, front grating, rear grating, and phase. The tunable laser driver circuit board sends a constant current to the gain section to maintain laser power. It converts the digital control signal from the FPGA into analog current outputs to the other 3 sections using high-speed DACs, thus controls the lasing wavelength. The requirements for the tunable laser driver and

control board are mainly imposed by the components after it, especially the wavelength converter, the arrayed waveguide grating (AWG; a static wavelength switch) and the optical network outside the optical router. The key requirements are stability in power and wavelength, accuracy in wavelength, and high speed in wavelength tuning.

Figure 22: Tunable laser and driver in the architecture

● The design of the tunable laser driver and control board: the first and second version

Figure 23 shows the block diagram of the tunable laser driver and control circuit. The D/A converters AD9731 control each current going into the front, rear or phase section. The input of AD9731 is the digital data of the current value. Because FPGA has limited output ports, it provides the most significant 4 bits of the data for front and rear currents. Dipswitches on the board control all the remaining bits and the phase current. Controlling only the most significant 4 bits is adequate to access the useful wavelengths for most of the experiments when channel count is low. Sometimes the wavelength is not accurate or unwanted side-lobes show up. Then the dipswitches will be adjusted to optimize the laser output.
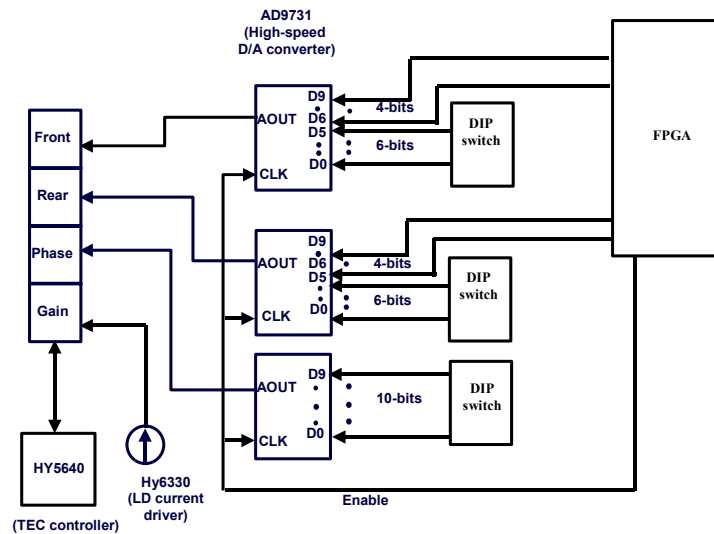
Figure 23:   Block diagram of the tunable laser driver and control circuit

Figure 24 shows the block diagram of the second version of the circuit. It introduces a table of the mapping from digital data to real current. For this purpose, an Static Random Access Memory (SRAM) and a micro-controller are introduced. This relieves the FPGA of tracking the current-wavelength relationships for individual tunable lasers. Instead each relationship is stored in the form of a look-up table and burned into the micro-controller. On start-up, the table will be loaded into the SRAM. Thus when FPGA sends the "code" for a particular wavelength, the corresponding digital current values will be searched and found in the SRAM. The DACs will then convert the digital value into analog current to tune the wavelength of the tunable laser. Also, an RS232 serial port is provided for the connection to a

computer. The computer can communicate with the micro-controller to perform more sophisticated tasks, for example automated tunable laser calibration.

Figure 24: Block diagram of the second version tunable laser driver and control circuit

According to the block diagram the circuit schematics were designed. The PCB layout and manufacturing were contracted to a PCB fabrication company. The assembly and testing of the board was done in our lab.

The wavelength switching characteristics of the tunable laser were tested. For static characteristics, 256 data points were measured completely to cover all combinations of the most significant 4 bits of both front and rear DAC. Figure 25 shows a picture of the second version tunable laser driver and control circuit board.

Figure 25: Picture of the second version tunable laser driver and control circuit board
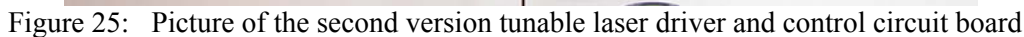
4.9 Deliverable 3c: Switch Fabric Integration Complete (Y3Q2)

The optical-label switching technology has made a key progress in providing low-latency and transparent switching desired for the next generation Internet. Achieving packet switching directly in the optical layer can facilitate convergence of optical networking and data networking without relying on redundant optical-to-electrical, and electrical-to-optical signal conversions.
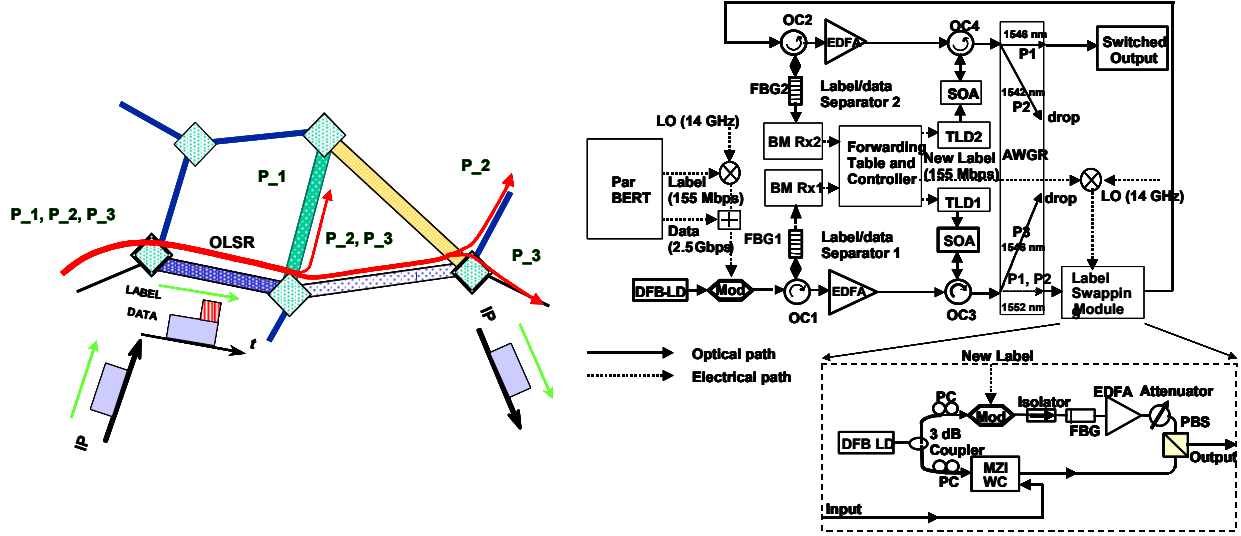
Figure 26: (a) A schematic of the emulated network, and (b) an experimental setup for the multi-hop emulation of this network. (Par BERT: Parallel Bit error rate Tester; PPG: Pattern Generator, LO: Local Oscillator; Mod: Modulator; FBG: Fiber Bragg Grating; BPF: Band pass Filter; OC: Optical Circulator; BM Rx: Burst Mode Receiver; SOA: Semiconductor Optical Amplifier; LO: Local Oscillator; MZI WC: Mach Zehnder Interferometer Wavelength Converter; TLD:Tunable Laser Diode; AWGR: Arrayed Waveguide Grating Router; PC: Polarization Controller; PBS: Polarization Beam Splitter)

The experiment emulates a network with multiple optical-label switching routers, each providing label-based packet forwarding. Figure 26 (a) shows an emulated optical-label switching network consisting of several optical-label switching routers (OLSR). Three types of packets, P1, P2, and P3 with labels L1, L2, and L3 respectively, ingress into the optical label switching networks. The first OLS router (OLSR1) performs the optical-label based forwarding of all three packets. OLSR1 forwards P3 north to a neighboring OLS router (OLSR2) and forwards P1 and P2 east to another neighboring OLS router (OLSR3). The OLSR3 in turn forwards P1 and P2 to two different output ports. For multi-hop scalable OLSR operations, data payload regeneration and label swapping/regeneration are beneficial. Figure 26 (b) shows the detailed structure of OLSR and setup for this network emulation. As Figure 26 (b) shows, the actual experiment places OLSR1 and OLSR3 form Figure 26 (a) on the same optical router system with multiple line cards and replaces OLSR2 with a drop port. The OLSR system consists of an optical-subcarrier multiplexing transmitter (SCM Tx), two optical label/data separators, two burst mode receivers (BMRx1 and BMRx2) for label detection, a field programmable gate array (FPGA) that implements the forwarding table and switching control, two tunable wavelength converters consisting of tunable lasers and semiconductor optical amplifiers (SOAs), a uniform-loss-cyclic frequency (ULCF) arrayed waveguide grating router (AWGR), a label rewriting module [5] and data receivers. The Parallel Bit Error Rate Tester/Pattern generator (ParBERT) synchronously generates the electrical label and payload signals. The LiNbO$_3$ external modulator modulates the continuous wave (CW) light from the distributed feedback laser diode (DFB LD) using a subcarrier multiplexed signal consisting of a baseband 2.5 Gb/s data payload and a 155 Mb/s label modulated onto a 14 GHz subcarrier. Hence, the modulated signal includes a double-sideband subcarrier label 14 GHz away from the center optical carrier frequency. The

combination of a fiber Bragg grating and an optical circulator achieves all-optical label extraction [6]. The BMRx asynchronously recovers the label contents from optical domain to electrical domain. The recovered label signal induces the forwarding decision inside the switch controller according to the routing algorithm in the FPGA. Based on the forwarding decision, the switch controller sends a control signal to the tunable laser (TLD) to switch to the designated wavelength [7]. The TLD generates a tunable probe light for the SOA1, which modulates the payload signals onto the new wavelength by cross-gain modulation. Payloads with different labels are converted onto different wavelengths corresponding to the desired output ports of the AWGR.
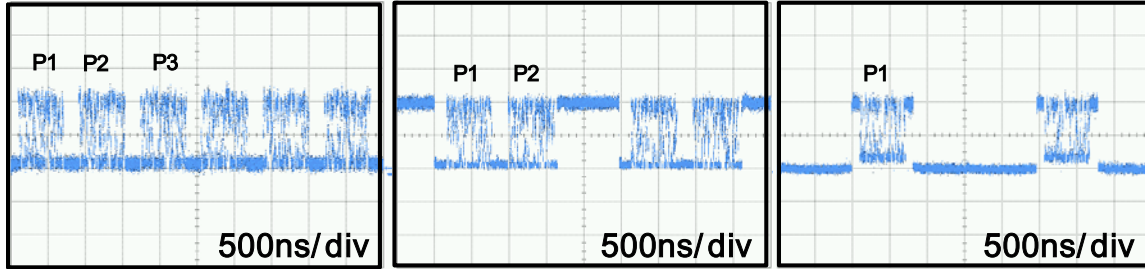
Figure 27: Scope traces showing (a) the incoming P1, P2 and P3, (b) P1 and P2 sent to the second hop after dropping P3, and (c) P1 at the final destination output after two-hop OLSR.

The ParBERT generates repeated patterns of packet 1 (P1), packet 2 (P2), and packet 3 (P3) with different labels (L1, L2, L3). Figure 27 (a) shows the three packets. The optical labels L1, L2 and L3 cause the wavelength of the TLD1 to be switched to λ 1 (1552nm), λ 1 (1552nm), and λ 2 (1546 nm) respectively according to the forwarding table. The tunable wavelength converter will convert P1 and P2 to wavelength 1552 nm and P3 to 1546 nm according to the optical-label based forwarding decision. After routing through the AWGR, P3 will be dropped and P1 and P2 go to the label-swapping module as shown in Figure 27 (b). This represents the OLSR1 in the emulated network. The switch controller generates new labels for payload P1 and P2 and drives the modulator inside the label-rewriting module with a 14 GHz carrier frequency. At the same time, payload P1 and P2 will be regenerated to the fixed wavelength (1555.7nm) in the SOA Mach-Zehnder Interferometer wavelength converter (SOA-MZI WC). The packets with the regenerated label and payload transmit to label / data separator 2. The BMRx2 recovers the new label contents and sends them to the switch controller. Again, according to the label contents, the switch controller sends control signals to the TLD2 to switch to the correct wavelengths. The new labels L1' and L2' cause the wavelength of the TLD2 to be switched to λ 1' (1546nm) and λ 2' (1542nm) respectively. TLD2 drives the SOA2 that converts the payload signal onto the desired wavelength by cross-gain modulation. P1 converted to 1546 nm will be routed to the destination port.   Figure 27 (c) shows P1 on the destination port. P2 converted to 1542 nm will be dropped after AWGR. The switched data payload P1 goes to the data receiver for BER measurements. The switching with new labels L1' and L2' emulates the OLSR3 in Fig. II.D.9.1 (a).

Packet by packet bit-error-rate measurements took place on the P1 at each hop. Figure 28 shows the measured data. Each packet is 600ns long with a 200ns guard time, thus each packet period is 800ns. The bit pattern was $2^{15}-1$ PRBS truncated into the packets.   The three curves in Figure 28 are for the optical baseband back-to-back and the payload signals after one and two hop OLSR, respectively. The signal after one hop shows about 0.7 dB power penalty compared to the baseband payload signal. However, a negative power penalty about 0.2 dB at BER=1e-9 appears after 2 hop OLSR, which is mainly due to the 2R regeneration in the SOA-based MZI WC and the decrease of the received average power after two packet-droppings.   The eye diagrams of the switched payload are shown as the insets in Figure 28. All eye diagrams show clear openings. The signal-to-noise ratio was higher for the second hop compared to the first hop due to the 2R regeneration of MZI WC. Also the XGM based SOA wavelength converters invert the logic of the signal which leads to the change of the average power of signal. This results in a

higher average power for the first hop. For the second hop the second XGM wavelength converter will invert the logic back to normal. For these reasons, the power received by the date receiver corresponds to different ratios of the real packet power for the 1-hop operations and the 2-hop operations. The combination of the 2R regeneration and the optical power change leads to the negative power penalty for the 2-hop operation.

Figure 28: BER test results of the cascaded OLSR (Insets: eye diagrams of the baseband payload and signals after OLSR)

5. TECHNICAL APPROACH AND RESULTS

5.1 Networking Architecture and Protocol
5.1.1 "Smart Edge Fast Core" optical packet switching network: architecture, scalability performance, and congestion control

5.1.1.1 Introduction
This report first presents a "Smart Edge Fast Core" architecture of optical packet switching (OPS) network, which achieve high-performance packet forwarding in the core, incorporating an intelligent traffic management at the edge. To handle diverse Internet traffic patterns, we exploit the availability of the electronic buffer at the edge to reshape the traffic profile. In particular, the performance enhancement schemes via the traffic-shaping function at the ingress edge and the redundant local drop ports at the egress edge are highlighted through simulation experiments. We compare the performance of the proposed OPS network with conventional IP router networks, and demonstrate the scalability advantage of OPS technologies.

To address the issue of congestion control, this report outlines an active congestion control architecture for the OPS network and presents the coordination mechanism among the core router, the edge router and

the NC&M (network control and management) system. Specifically, we elaborate on two main components of this architecture: the link–utilization-based congestion detection and the transmission rate adjustor for TCP acknowledgement (ACK) segments. Simulation experiments examine their efficacy for network performance improvement.

5.1.1.2 "Smart Edge Fast Core" OPS network architecture

The proposed network architecture includes OLS-based packet switching routers at the core with smart electronic routers at the edge. The core network has an all-optical data plane. The core OPS router consists of arrayed-wavelength-grating-routers, tunable wavelength converters, fixed wavelength converters, and a switch controller. The OLS router forwards packets according to the content of the forwarding table in the switch controller. When contention occurs in the switch fabric, it will first attempt to resolve contention by forwarding the packet on an alternative wavelength. If none is available, it will attempt to forward the packets to the fiber delay lines (FDL) for time buffering. If there is no free FDL, it will deflect the packet to a secondary preferred output port. In contrast to conventional electronic routers with contention resolution primarily in the time domain, the additional wavelength domain provides appealing potential to obtain high-performance all-optical switching capacities [1]. In the control plane, the optical router taps off the optical-label from the attached packet, and makes the forwarding decision based on the label content and the forwarding table. The optical label contains routing and control information and is carried in-band by sub-carrier multiplexing to achieve encoding within the same optical wavelength band [2].

Due to the adverse influences of bursty Internet traffic, the optical contention resolution schemes alone cannot sufficiently maintain a reasonable performance under a high traffic load. Meanwhile, statistical traffic data show that nearly half of the IP packets are 40 to 52 bytes in length and the IP packet size follows a unique distribution with peaks at 40, 576, and 1500 bytes [3]. To efficiently handle these diverse traffic patterns, the electronic memories at the ingress edge routers are exploited to achieve traffic-shaping function by assembling "jumbo optical packets" from client IP packets of the same destination and of common attributions. The creation of an optical packet is triggered by a maximum payload size (MPS). Whenever the buffer occupancy of the assembly queue reaches the MPS, an optical packet will be assembled. To avoid excessive queuing delay, the assembly mechanism adopts a timeout period. The expiration of the assembly timer also generates an optical packet. This packet-assembly mechanism can smoothen the optical packet size distribution and reduce the burstiness of Internet traffic, thereby resulting in a notably lower packet-loss rate [4]. At the egress edge, optical packets are forwarded from the core router through the local drop ports. The limited number of the drop ports poses a constraint to the transmission capacity from the optical core to the electronic edge. To avoid over-dimensioning optical resources at the core, it would be beneficial if optical packets would promptly reach edge routers. From this perspective, we propose to add redundant drop ports to gain further performance improvement. For a given number of clients (N) we will consider cases with $2N$ drop ports, thus requiring an Nx$2N$ egress edge router.

To examine the efficacy of these schemes, we use the OPNET modeler to conduct simulations based on a six-node network topology (Figure 29 (a)) with a representative nodal architecture (Figure 29 (b)). The baseline network has edge routers with four add/drop ports without traffic shaper (thereafter denoted as "4×4 edge, no shaper"). We also simulate the edge router with four add ports and eight drop ports (4×8 edge); note that it includes redundant drop ports. Each link operates in the WDM mode with four wavelengths transmitting at 2.5 Gbps. Each node is capable of performing wavelength conversion, optical buffering, and deflection routing. The holding time of the FDL is equal to the maximum optical packet length, and the number of FDL's is set at K-1 (here K is the nodal degree of the optical router). The maximum payload size (MPS) is set at 9000 bytes for traffic shaping function. The simulations are driven by self-similar traffic sources with Hurst parameter 0.8. Figure 30 shows the IP packet length distribution used in this work, which has been proved to accurately reflect the spikes of packet lengths seen in the Internet [5]. Figure 31 shows the network-wide packet-loss rates plotted against the

average offered transmitter load. It is clear that traffic shaping together with the redundant local drop ports result in noticeably smaller PLR's, compared with the baseline network without the support of a smart edge. For instance, when load=0.5, the PLR in the baseline network is 0.883%, which reduces to 0.229% with the help of the redundant local drop ports alone, and further down to 0.0226% with the traffic-shaping function.

(a)   (b)
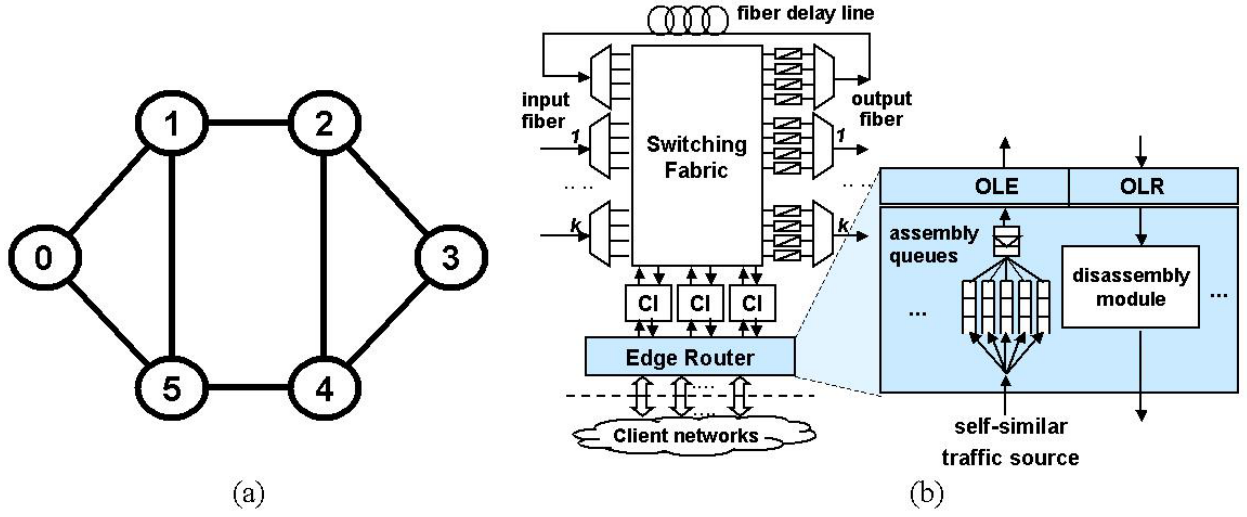
Figure 29: (a) The network topology and (b) the nodal architecture used for simulation in this report
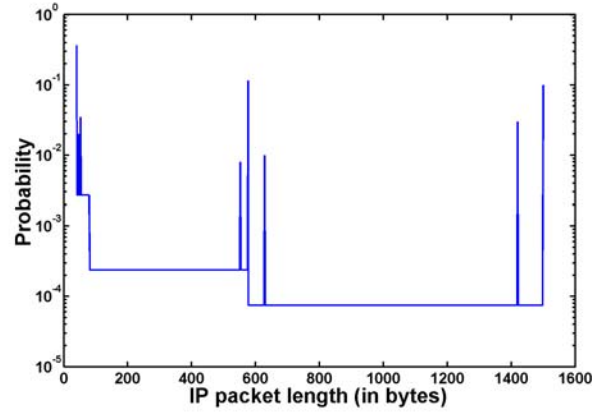
Figure 30: IP packet length distribution

29
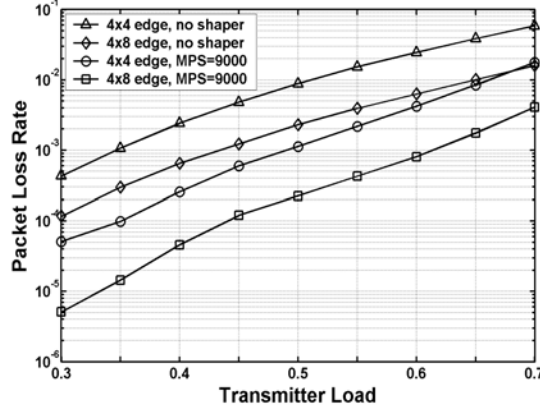
Figure 31: Packet-loss rates with various network configurations

5.1.1.3 Scalability performance: optical packet switching network vs. electronic packet switching network

To highlight the scalability advantage of the OPS network, this section compares its performance with the electronic IP router network under the same transmission capacity, traffic load and network topology, as described in Section 5.1.1.2. For the electronic network, we use four OC-48 parallel links to connect the adjacent IP routers, instead of one WDM fiber in an OPS network. To manage these parallel links, the electronic network runs the OSPF protocol to enable the Equal-Cost Multi-Path (ECMP) support on the IP routers.

In an IP network with ECMP function, the overall network performance heavily depends on the load balancing schemes across the multiple parallel paths. Currently IP switching supports two general load balancing modes: per-packet and per-destination [6]. Per-packet load balancing allows the router to send successive data packets over multiple paths without regard to individual hosts or user sessions. Per-packet load balancing ensures balancing over multiple paths, but packets for a given host pair might take different paths which could cause reordering of packets. In contrast, with per-destination load balancing, packets for a given destination are guaranteed to take the same path even if multiple paths are available, but it may result in unequal load sharing among multiple paths. Our simulations explore their effects on the network performance. Each electronic IP core router is fed with a self-similar traffic source to generate IP packets following the realistic packet-length distribution (as described in Section 5.1.1.2); the generated traffic is equally dispersed to the other core routers. Each linecard in the IP router provides one OC-48 port and contains a 2Mbyte buffer.

As Figure 32 demonstrates, the per-destination load balancing scheme results in a noticeably high packet-loss rate in the simulated network, where we note that the neighbor IP routers connect with multiple parallel links. In a network with a high traffic load, a per-destination scheme would cause several links into congestion, mainly because it forbids IP routers to forward packets through other available links to their neighbors. Since some links become overloaded, the queue would build-up quickly in the IP routers and thus, the sustained packet-loss would occur inevitable. Such undesirable link-saturated scenario cannot be avoided by solely increasing the depth of the memory buffer, primarily because the conventional wavelength-blind IP router is incapable of employing effective load balancing in the wavelength domain. This finding provides an explanation to the "link-bundle" problem observed by recent network tests on high-end electronic routers [7]. As reported in [7], enabling a link-bundle feature of IP routers degrades throughput and latency, because the algorithms implemented by router vendors is incapable of efficiently supporting parallel links between the adjacent routers. The simulation results of the OPS network indicate that the wavelength-awareness of the OPS technologies enables a scalable

solution to efficiently utilizing the multi-path offered by DWDM. The combination of contention resolution schemes at the core and performance enhancement mechanisms at the edge helps the OPS network to achieve a good performance to meet the requirements for the next generation IP/DWDM networks.
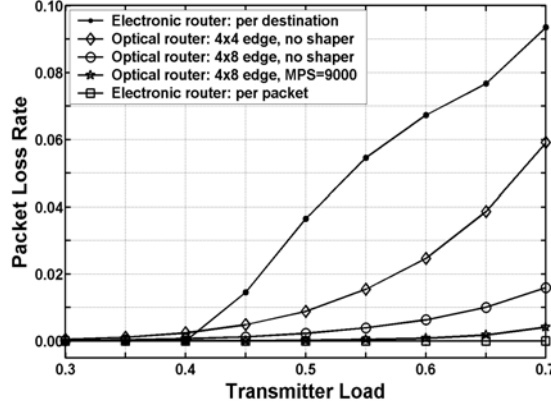
Figure 32:   Packet-loss rate comparison: optical vs. electronic networks

5.1.1.4 Active Congestion control architecture for OPS networks

The goal of congestion control is to maximize the throughput and to minimize the packet-loss rate. Recently, active congestion control has drawn significant attention in electronic packet-switched networks. Some pertinent methods, such as active queue management (AQM) and explicit congestion notification (ECN), have been extensively investigated [8,9]. Although the same rationales behind these schemes apply to the optical packet-switched (OPS) networks, the conventional queue-based solutions are unsuitable for the OPS domain primarily because the OPS network lacks the memory buffer to deploy the "store-and-forward" queuing system. Thus, it is essential to take a new view for these issues by considering the key features of the OPS networks.

In light of the observation that 80%-95% of Internet traffic uses TCP [3], our proposed active congestion control is of TCP-awareness. The key idea is to detect incipient network congestion and convey congestion notifications early enough to TCP sources, so that the sources are able to throttle down their transmission rates before any sustained packet-losses occur.

In contrast to conventional schemes that rely on queue occupancy to detect congestion, the proposed scheme infers the network congestion status directly from the measured link-utilization. The NC&M agent embedded in each optical router continuously monitors the traffic rate. For each fiber link, the traffic rate r is estimated by using an exponential averaging: $r_{new} = (1-e^{-T/K}) \times S/T + e^{-T/K} \times r_{old}$, where T is the inter-arrival time between the current and the previous packet, S is the size of the arrival packet, and K is a constant. This traffic rate estimator proved to be converged to the actual rate under a wide range of network conditions [10]. Since OPS routers are capable of balancing traffic load among multiple wavelengths within the same WDM link, the link-utilization could be approximated by only measuring one of those wavelengths, thereby reducing the computing burden and making it more feasible. With a predefined maximum utilization threshold MAX_{th}, the NC&M detects a link to be congested if $r \geq C \times MAX_{th}$ (C is the link rate) at all times during an interval of length d. The proper choice of the MAX_{th} value will target sustaining of the link traffic load below the 0.6~0.7 level [11]. Once the NC&M detects the incipient congestion, it sends out a congestion notification to the edge router, which is responsible for providing congestion control feedback to the TCP source, as described below.

In TCP congestion control, ACK-clocking serves as a regulator to clock out the transmission of new data from TCP sources. TCP data transfer in the forward path heavily depends on the transmission rate of ACK segments over the reverse path. From this perspective, our proposed scheme employs an ACK

transmission rate adjustor at the edge router, where a separate queue for the reverse ACK segments is created by using the available electronic memory. The edge router provides fast feedback to TCP sources by dynamically adjusting the ACK transmission rate based on current network status. After receiving the congestion notification from the NC&M system, the edge router initiates congestion control by decreasing the ACK transfer rate $r_a=r_N/\alpha$, where $\alpha>1$ is the decrease factor, and r_N is the ACK transmission rate in the case $(C\times MIN_{th})<r<(C\times MAX_{th})$, C is the link rate. Here MIN_{th}, the minimum utilization threshold, is adopted to sense the link underloaded status, which triggers the edge router to increase the ACK transmission rate by $r_a=r_N\times\beta$ with the increase factor $\beta>1$. The detailed description of this mechanism is as follows:

if ($r\geq C\times MAX_{th}$) for a time-interval of length d

$$r_a=r_N/\alpha\,;$$

if ($C\times MIN_{th}<r<C\times MAX_{th}$) for a time-interval of length d

$$r_a=r_N\,;$$

if ($r\leq C\times MIN_{th}$) for a time-interval of length d

$$r_a=r_N\times\beta\,.$$

Compared with the conventional end-to-end solution like ECN, the proposed scheme promotes congestion avoidance at the very early stage and provides the network operator predictable resource management via the tunable parameters. Rather than dropping packets to indicate network congestion in conventional AQM approaches, the ACK rate adjustor can implicitly signal the TCP source with a shortened control loop, which reduces the reaction time and results in a higher throughput. Since there is no modification of the end-system TCP, the proposed scheme preserves the end-to-end semantics of TCP congestion control, which facilitates its deployment to integrate with the legacy networks.

5.1.1.5 Performance simulation of the active congestion control mechanism

To explore the feasibility of the proposed scheme in OPS networks, we conduct simulation on the OPNET modeler by considering a six-node topology (Figure 33). Twelve TCP connections are established between a set of senders and receivers through the core OPS network. Fig. II.E.1.a.1(b) shows the architecture of the core router, which is capable of performing wavelength conversion, optical buffering, and deflection routing. In the core network, each fiber link operates in the WDM mode with four wavelengths transmitting at 2.5 Gbps. The link rate between the TCP workstation and the core network is 1Gbps. This work uses the TCP *new Reno* (containing slow start, congestion avoidance, fast retransmit and fast recovery) to generate TCP data traffic through greedy bulk-data transfers. The maximum segment size is 1460 bytes, and the maximum congestion window is 4 Mbytes with the window scaling option. To focus on the data traffic over the forward path from node 0 to node 3, we assume the reverse path for ACK transfer is loss-free. In addition to explicit TCP traffic, we use self-similar traffic sources to generate background UDP traffic. For nodes 1, 2, 4 and 5, each one connects with four self-similar traffic sources via local add ports to generate UDP traffic with transmission load 0.1. The generated traffic is equally dispersed to the other core nodes. Note that Fig. II.E.1.a.5 does not show those UDP sources for the purpose of simplicity. At the node 0, the client interface implements the aforementioned ACK transmission rate adjustor, where r_N is set to 10^6 ACK segments per seconds, $\alpha=\beta$ $=2$. To configure the congestion detection algorithm in the NC&M system, we set $MAX_{th}=0.65$, $MIN_{th}=0.40$, and $K=D=200$ μs.
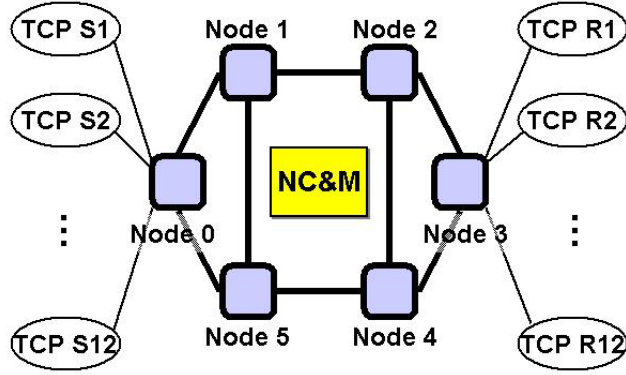
Figure 33:   Network topology used in simulation

Due to the explicit TCP traffic and underground UDP traffic, the link from node 1 to 2 will become overloaded at some time-points during the simulation period. Figure 30 and 31 depict the measured link-utilization over the link 1->2 in a network with and without congestion control, respectively. By dynamically adjusting the ACK transmission rate based on network status, the proposed scheme results in a smoother traffic load (Figure 34), and prevents the adverse load oscillations that are observed in the network without active congestion control (Figure 35).
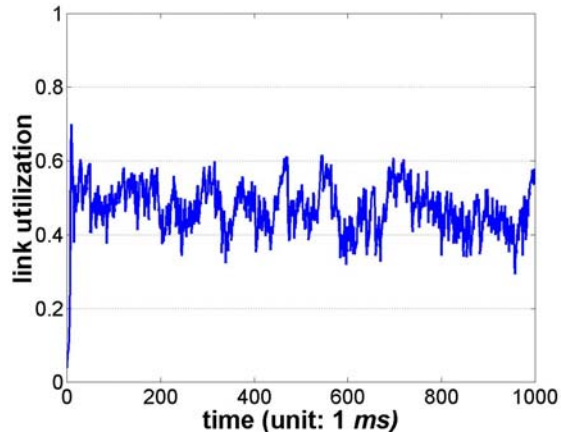
Figure 34:   Utilization of link 1->2 in the network with active congestion control
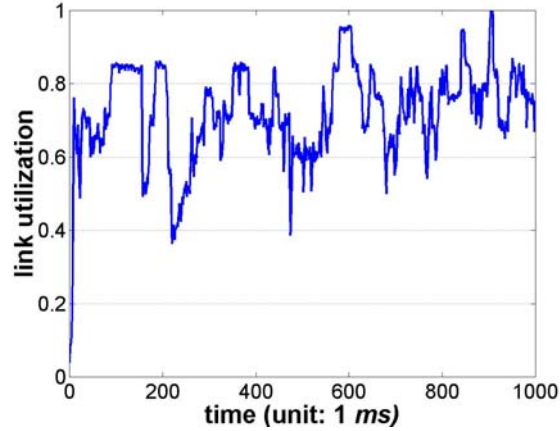
Figure 35:   Utilization of link 1->2 in the network without congestion control.

To evaluate the performance improvement, we use the metrics of the network-wide packet-loss rate, the average TCP goodput, and the fairness index. The TCP goodput measures the efficient throughput by excluding the unnecessary retransmissions. We measure the fairness by using fairness index f defined in [12]: if there are n concurrent connections and the throughput achieved by connection i is x_i, then f is the ratio between $[SUM(x_i)]^2$ and $n \times [SUM(x_i^2)]$. Table 2 summarizes the simulation results. Compared with the "no congestion control" network, the active congestion control dramatically reduces the packet-loss rate by one order of magnitude. As expected, such a lower packet-loss rate avoids the unnecessary retransmission and time-out in TCP transfers, and leads to a higher goodput. It also shows that the proposed scheme attains a better fairness in terms of the bandwidth allocation among the competing TCP connections.

	no congestion control	active congestion control
average utilization of link 1→2	0.719	0.4685
packet-loss rate	0.0442	0.00130
average goodput	25.5 Mbytes/sec	38.4 Mbytes/sec
fairness index	0.918	0.985

Table 2 Simulation results.

This section presented a "smart edge fast core" OPS network architecture, which includes OLS-based packet routers at the core and performance enhancement schemes at the edge. Both the traffic-shaping function (at ingress edge) and the redundant local drop ports (at egress edge) proved helpful to relieve the burden on the optical contention resolution in the core, thereby significantly reducing the packet-loss rate. Based on the proposed architecture, we investigated the scalability performance of OPS networks. The simulation results indicate that the OPS technology provides a scalable solution for a seamless integration of the electronic IP/MPLS layer and the optical DWDM layer.

To achieve the congestion control in OPS networks, we proposed a TCP-aware active congestion control architecture, and presented the coordination mechanism among the core router, the edge router, as well as the NC&M system, which brings a new vision to achieve high-performance in the OPS domain. Simulation experiments examined the performance improvement achieved by the proposed scheme. Preliminary results demonstrated that the link-utilization-based congestion detection and the ACK transmission rate adjustor could dramatically reduce the packet-loss rate, thereby efficiently improving the network throughput. Since the proposed scheme shortens the control loop by reacting to the incipient

congestion at the network edge, it is expected that the performance improvements will become even more significant as the network scales. Based on the outlined control architecture, we are currently exploring the optimum solution by fine-tuning the system parameters.

References

[1] S.Yao, et al., "All-optical packet-switched networks: a study of contention resolution schemes in an irregular mesh network with variable-sized packets," in Proc. SPIE OptiComm'00, Richardson, Oct. 2000.
[2] W. Shieh and A.E. Willner, "Demonstration of output-port contention resolution in a 2*2 WDM switching node based on all-optical wavelength shifting and subcarrier-multiplexed routing-control headers," OFC '96 (1996).
[3] K. Claffy, et al., , The nature of beast: Recent traffic measurements from an Internet backbone. http://ipn.nlanr.net/Papers /Inet98/ index.html.
[4] F. Xue, et al., "The performance improvement in optical packet-switched networks by traffic shaping of self-similar traffic," in OFC'02, Anaheim, Mar. 2002.
[5]Agilent Technologies, Insight: for testing IP/optical networks: Edition 1. http: //advanced. comms. agilent.com /insight /2001-08/index.htm.
[6] Cisco Systems, Troubleshooting Load Balancing Over Parallel Links Using Cisco Express Forwarding (CEF). http:// www.cisco.com/warp /public/105/ loadbal_cef.html .
[7] David Newman, The Trouble with Trunking. http:// www.nwfusion.com/reviews/2001/0416rev2.html
[8] S. Floyd, and V. Jacobson, "Random early detection gateways for congestion avoidance," IEEE/ACM Trans. Networking, vol.1, 397-413 (1993).
[9] K. Ramakrishnan, and S. Floyd, "A proposal to add explicit congestion notification (ECN) to IP," RFC 2481, Jan. (1999).
[10] I. Stocia, S.Shenker, and H. Zhang," Core-stateless fair queueing: achieving approximately fair bandwidth allocations in high speed networks", Sigcomm'98, (1998).
[11] L. Tancevski *et al*., "Optical routing of asynchronous, variable length packets," *IEEE J. Select. Areas Commun.*, vol.18, 2084-2093 (2000).
[12] D. Chiu, and R. Jain, "Analysis of increase and decrease algorithms for congestion avoidance in computer networks*," Computer Networks and ISDN Systems*, vol. 17, 1-14 (1989).

5.1.2 GMPLS traffic engineering for optical label-switching networks

5.1.2.1 Background

In this research work the purpose is to evaluate the influence of congestion on network performance and design a traffic engineering method to alleviate congestion and improve network performance. Opnet models were built to simulate the effect of congestion and traffic engineering method. Previous GMPLS traffic engineering work assumed that packet switching is performed only on an electrical layer such as IP or ATM leaving no possibility for optical-packet switching. This work extends GMPLS traffic engineering in a multi-granularity optical label switching network, where both packet and wavelength light-path (LP) switching take place directly in the optical layer. In a GMPLS context, this work refers to the first kind of label switched path (LSP) as "optical-packet switched LSP", the second kind "light-path LSP". The light-path LSP is used as traffic engineering tunnel. The optical-label switching router (OLSR) introduced in this work supports both packet switching and optical cross connection (OXC). The control plane uses the Generalized Multi-protocol Label Switching (GMPLS) protocol with new extensions. In real life data networks, sustained large volume traffic often occurs between two nodes (referred to in this work as a "traffic surge'), such as the daily synchronization of commercial databases. Simulation results show that the dynamic adaptive traffic-engineering algorithm proposed in this work can effectively

alleviate network congestion problems caused by such traffic surges over the optical-label switching network.

5.1.2.2 Optical label switching router (OLSR) architecture

The switching fabric part of the optical-label switching router is shown in Fig. II.E.1.b.1. Its architecture supports both packet switching and circuit switching. The OLSR switching fabric consists of tunable wavelength converters on the input side, an arrayed-waveguide-grating-router (AWGR) [7] in the middle, fixed wavelength converters on the output side, and a switch controller with a forwarding table. The AWGR can perform wavelength dependent routing via the well-known routing function. The tunable wavelength converters tune their conversion output wavelength to an appropriate wavelength so that the AWGR can forward the packet to a desired output port. The fixed wavelength converter then converts the output wavelength to the one that is proper for transmission through the fiber. The OLSR is capable of resolving contention in wavelength-time-space domain to reduce the packet loss rate [4]. When the OLSR switches a light-path circuit, it will tune the input wavelength to the desired wavelength and keep the wavelength until the circuit is torn down. In this case the AWGR performs OXC functions. The architecture of the OLSR inherently supports both optical packet and light-path circuit switching.
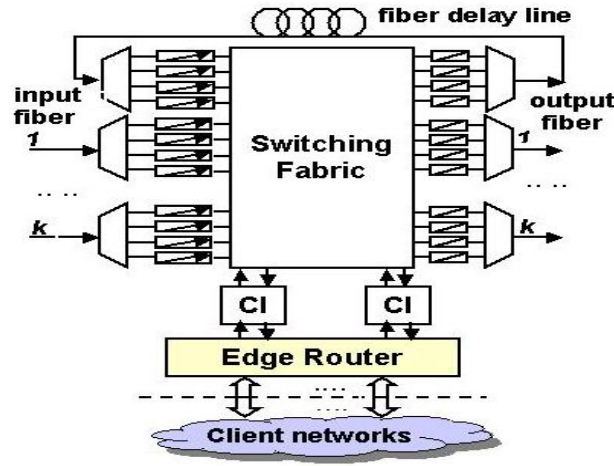
Figure 36:   Schematic of OLSR node architecture illustrating the optical switch fabric

5.1.2.3 Network architecture and simulation parameters

Simulation is based on the network architecture shown in Figure 37. Each node has the architecture introduced in Section 5.1.2.2. Each link is a fiber that consists of four wavelength channels each running at 2.5GB/s. The OLSR balances the wavelength channel load in a fiber link by performing "round robin" load balancing. To limit the scale of the problem, this work considers the situation that the traffic surge occurs only between node 0 and node 3. 1000 client stations are attached to node 0, and these clients communicate with an equal number of clients attached to node 3. The load of this traffic on the access wavelength channel is denoted by ρ_o. The network is also loaded with background traffic that is uniformly distributed over the network, the load of which is denoted by ρ_b. Background traffic and traffic surges are both generated by self-similar packet sources. The Hurst parameter is 0.8. The packet size distribution is modeled after the measured actual distribution, with a maximum packet size of 1500 bytes.
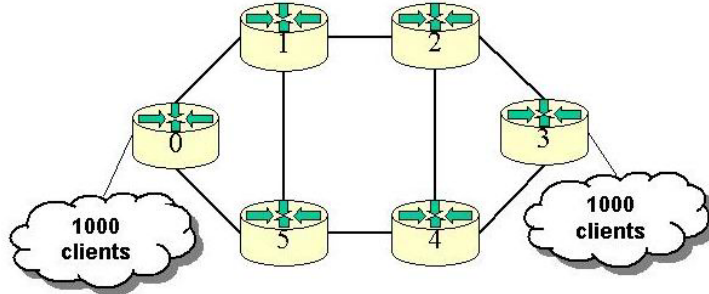
Figure 37:   The optical-label switching network architecture used for simulation

5.1.2.4 Traffic engineering and GMPLS control plane

When a traffic surge occurs between two edge nodes, it may cause congestion along its path. And if the wavelength-time-space domain contention resolution fails to resolve the packet contention, the OLSR will drop the packet. The proposed traffic engineering method is intended to reduce contention and packet drop by provisioning a dedicated light-path for the traffic surge, that is to put the surge traffic in a light-path traffic engineering tunnel. But doing this will also reduce the available wavelength resource for packet contention resolution. Hence, provisioning a dedicated light-path will not always reduce PLR. This work proposes a dynamic traffic engineering method, investigates the relationship between PLR and traffic load conditions, and makes the traffic engineering process adaptive to the load conditions by adjusting a set of thresholds, so that it ensures that provisioning a dedicated light-path will always reduce PLR. The network control and management system monitors the traffic load by the statistical data collected by the OLSR and the edge routers. If the traffic between an ingress and egress node pair meets some pre-set threshold conditions and lasts longer than a pre-set duration, the traffic engineering process will provision a light-path LSP between the ingress and the egress node. Several factors must be considered in this decision-making: the available resource of the network, the load of the traffic surge, and the load of the links over which the light-path will be provisioned. The light-path is dynamic. The thresholds are also dynamic values dependent on these factors. By adjusting the thresholds according to the traffic and resource conditions, the traffic engineering process is adaptive. The GMPLS control plane is used for the light-path LSP set-up.

The previously described traffic engineering process involves moving the flows carried by the packet switched LSPs to a circuit switched light-path LSP. To ensure successful set-up, the light-path LSP being provisioned should share the resource with all of the packet-switched LSPs it is intended to replace. So the light-path LSP should be in a shared resource group that includes itself and all of the packet switched LSPs.

5.1.2.5 Simulation results

Two kinds of PLRs are considered here: packet loss rate when a light-path is provisioned between node 0 and node 3 (PLR-LP), and packet loss rate when there is no light-path between node 0 and node 3 (PLR-PK). PLR-LP indicates network congestion after traffic engineering, while PLR-PK indicates that before traffic engineering. These two kinds of PLRs show the effect of dynamically provisioned light-path under different traffic load conditions. Fig. II.E.1.b.3(b) shows that at each background traffic load level, there exists a threshold of traffic surge load, if the traffic surge load exceeds the threshold value, PLR-LP will be lower than PLR-PK. The threshold is the point where the PLR-LP and PLR-PK curves cross. Because PLR-LP is not related to surge load, the PLR-LP curves shown in Figure 38 (b) are horizontal lines. The value of PLR-LP at each background load level is derived from Figure 38 (a).

Simulation results show that a traffic surge in a wavelength channel can substantially increase the packet loss rate and dynamically provisioned light-path can effectively reduce PLR. For example, in the case the

background traffic load ρ_b = 0.3: without traffic surge, the PLR-PK is about 0.0007. If there is a traffic surge between node 0 and node 3 with load ρ_o = 0.7 in only one wavelength channel, the PLR-PK goes up to 0.008. The PLR-PK is 10 times higher, although the average network load is increased only by about 0.005, because the traffic surge causes congestion in its path and thus increases contention. But in the same condition when ρ_b = 0.3 and ρ_o = 0.7, PLR-LP is about 0.0007 (shown in Figure 38 (b)). So PLR-LP is 10 times lower than PLR-PK. This result shows that dynamic traffic engineering method proposed in this work can effectively reduce PLR when the background traffic load is relatively low. In this situation, traffic surge load is always higher than the threshold; hence provisioning a light-path for the traffic surge can always improve the network performance.

If the background traffic load is relatively high, the threshold for dynamic light-path provisioning increases. When ρ_o is less than the threshold, provisioning a light-path for the traffic surge will increase PLR. Take the example of ρ_b = 0.5. If ρ_o = 0.96, PLR-PK (0.022) is 2.5 times PLR-LP (0.0085). But if ρ_o = 0.2, PLR-PK (0.0061) is less than PLR-LP (0.0085). Figure 38 (b) shows that, for ρ_b = 0.5, the threshold for dynamic light-path provisioning is ρ_t = 0.38. Table 3 shows the thresholds corresponding to different background loads.
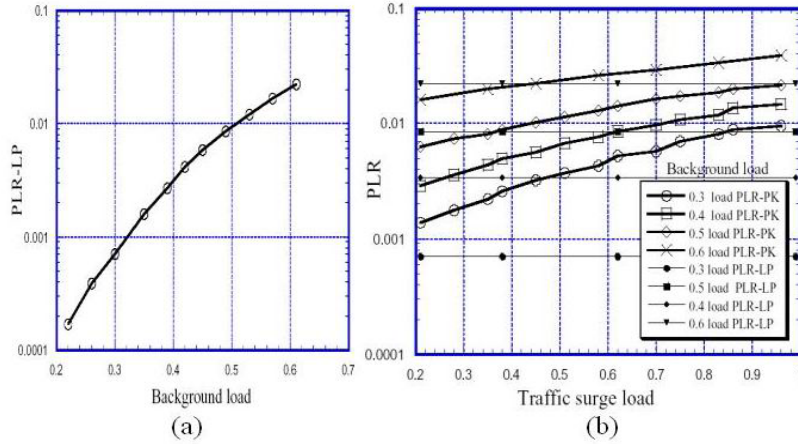
(a) PLR-LP vs Background load   (b) PLR-LP and PLR-PK vs traffic surge load

Figure 38: PLR under various load conditions

Table 3   Light-path provisioning threshold

ρ_b	0.3	0.4	0.5	0.6
Threshold ρ_{th}	~0.1	0.26	0.38	0.45

5.1.3 Optical Switching Fabrics, Paradigms, Survey and Technology Studies

5.1.3.1 Introduction

The network of the future is heralded as being all optical, data-rate independent and end-to-end compatible. Optical switches will be firmly entrenched in this vision. As networks continue to mature from electrical-based to all optical, different optical switching paradigms will address different demands and coexist together. Switch architectures, technologies, protocols, and control plane designs play an important role in the next generation optical networks to provide efficient contention-free optical packet buffering, switching, and scheduling with QoS/CoS guarantee. In this report, we further investigate the performance of such node architecture by simulation and illustrate the dominant factor affecting the packet loss rate in the architecture. We incorporate some new schemes, i.e., electronic circulation buffer,

wavelength domain service differentiation, and electronic domain buffer management schemes, to rich the contention-resolution approaches. Simulation results show that such extended schemes can efficiently support the DiffServ model.

5.1.3.2 Contention resolution schemes of optical switching fabric

In this section, at first we will investigate the proposed contention resolution schemes, then we will investigate new contention resolutions which may rich current schemes and improve the performance of the optical networks. Three contention resolution schemes have been proposed in the optical switch architecture based on feedback fiber delay lines (Figure 39). The contention of packet transmission can be resolved in three domains: wavelength, time, and space. When a packet comes in the node, the controller will check whether there's an available free wavelength in the destination output port. If there are no available free wavelength in the destination port, the controller will check whether there are some free wavelength buffer delay lines to buffer this packet. If the above two methods fail, the packet that loses a contention will attempt an alternate preferred output port. If this scheme still fails, the packet will be dropped.
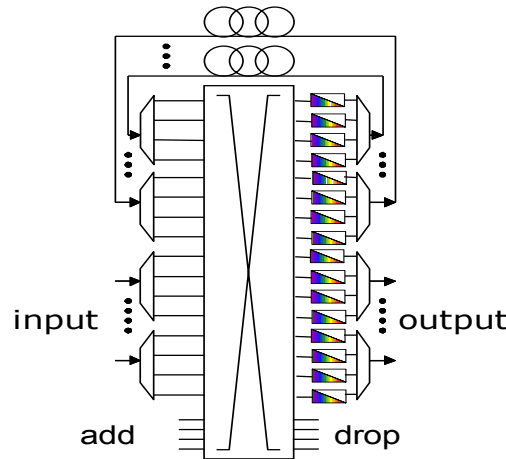
Figure 39:   Single-stage optical switch architecture based on feedback fiber delay lines.

Based on the above work, we will investigate the performance of such switch architecture by simulation and illustrate the dominant factor affecting the packet loss rate in this architecture. Also we will incorporate a new contention resolution scheme, i.e., electronic circulation buffer, to rich the contention-resolution approaches. The proposed new architecture is shown in Figure 40. The performance will be investigated by simulation.
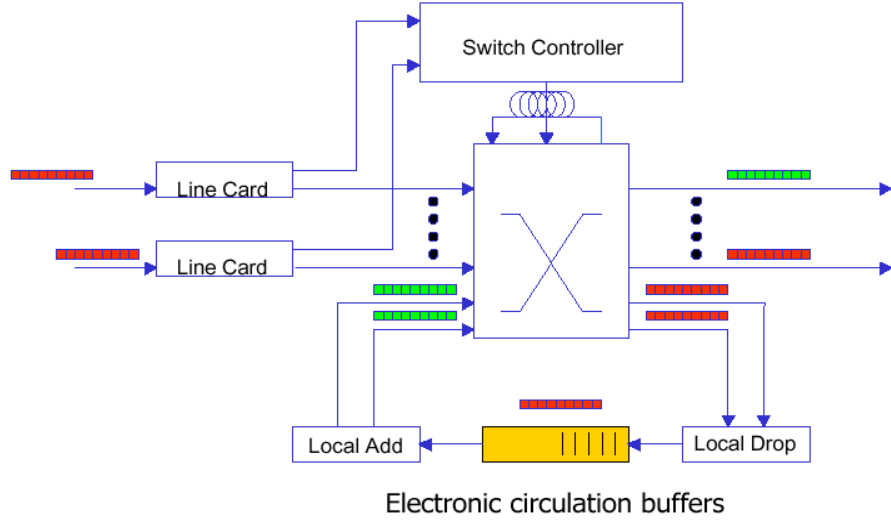
Figure 40:   Proposed switch architecture with electronic circulation buffer.

a.   *Simulation network configuration*

Figure 41 shows the 6 node-4 wavelength-2.5 Gps network topology we've chosen. Each node denotes an OLS router. Each WDM link contains 4 wavelengths, and each wavelength operates at 2.5Gb/s. The aggregate capacity for one link is 10 Gb/s. There're four optical delay lines in each node which equals the maximum number of interface ports. There're six traffic resources and each can inject 10 Gb/s traffic into the optical network.
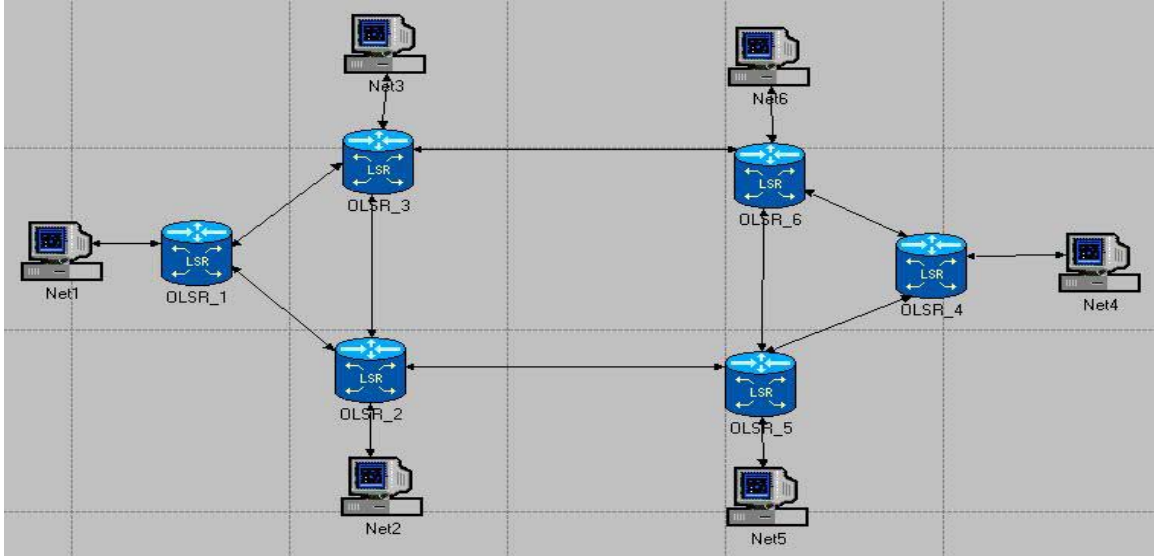
Figure 41:   Simulation Network Topology

b.   *Packet loss rate analysis*

Optical packets may be dropped at both the core optical nodes (intermediate optical switching nodes) and edge optical nodes (destination optical switching nodes). At the intermediate nodes, if the optical packet can not find the available wavelengths, fiber delay lines, and deflection routes, then it will be dropped at the transmission interfaces. PLR_TI is used to note the packet losses at the core optical nodes. When the

optical packet arrives at the destination optical nodes, if the packet can not find the available local receivers, then the packet will be dropped at the edge client interfaces. PLR_CI is used to note the packet losses at the edge optical nodes. Figure 42 and Figure 43 show the comparison of packet loss rate PLR_TI and PLR_CI with the traffic load of 0.3 and 0.8 respectively. The simulation data shows that PLR_CI is about as 100 times as PLR_TI which means most of the packets are dropped at the destination nodes where the packets can not find free local receivers.
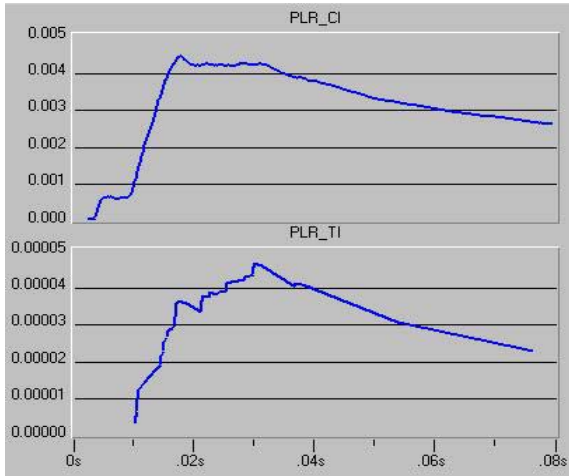
Figure 42:   PLR_CI and PLR_TI (Load 0.3)   Figure 43:   PLR_CI and PLR_TI (Load 0.8)
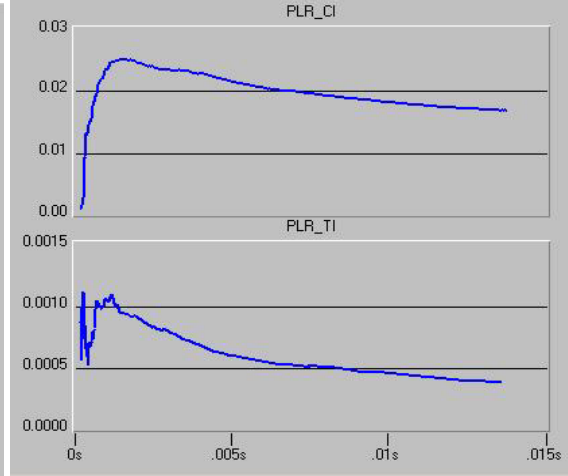
c.   Decrease packet loss rate PLR_TI at core nodes with electronic circulation buffers

Traditionally the contention of packet transmission in optical switching nodes can be resolved in three domains: wavelength, time, and space. When a packet comes in the node, according the destination address of the packet, the optical router controller checks the forwarding table and decides the output port for the packet. Then the controller will check whether there's an available free wavelength in the destination output port. If there's no available free wavelength in the destination port, the controller will check whether there's some free wavelength buffer delay lines to buffer this packet. If the above two methods fail, the packet that loses a contention will attempt an alternate preferred output port, which is determined by a pre-computed deflection-routing table based on some routing algorithms, such as static shortest-path routing. If all of the above three ways fail, the packet will be dropped directly. However, electronic circulation buffers (which may be implemented in edge routers) may be the fourth candidate to solve the contention resolution in optical switches. The basic idea behind is that if the first three ways fail, the packet will attempt an available local receiver and then will be sent to the electronic buffers via this receiver for re-transmission instead of dropping it directly.

Figure 44 shows the impact of electronic circulation buffers on the core nodes packet loss rate PLR_TI. The result shows that ECB may decrease PLR_TI greatly.
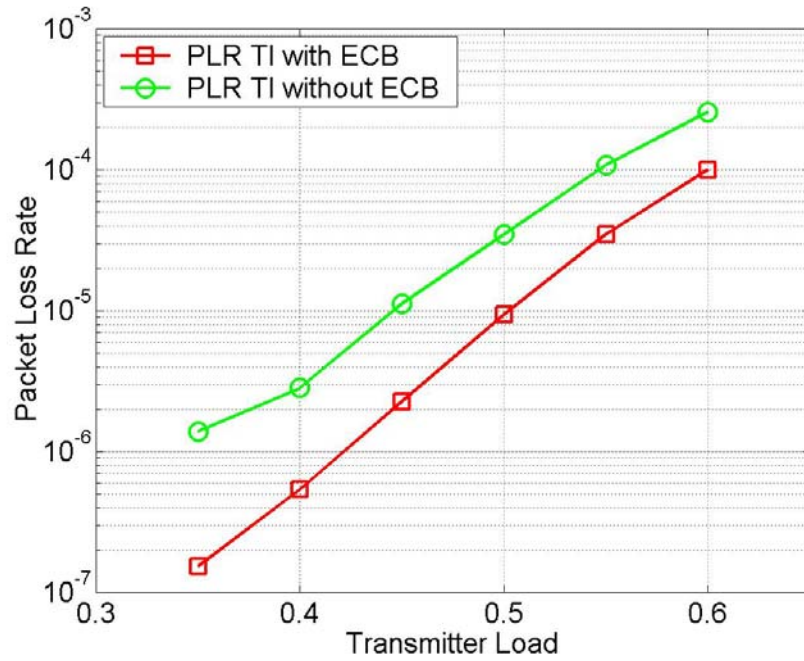
Figure 44:   Impact of ECB on the core nodes packet loss rate PLR_TI

d.   *Decrease packet loss rate PLR_CI at edge nodes with more local receivers*

Figure 45   shows the impact of the number of local receivers on the edge nodes packet loss rate PLR_CI. The result shows that a much lower PLR_CI is achieved by doubling the number of the local receivers N (N=4).
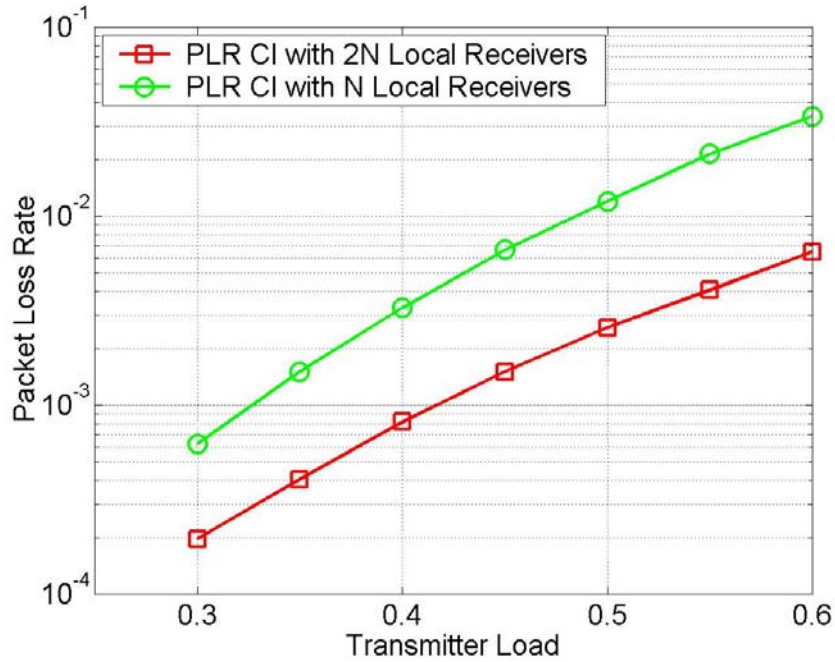
Figure 45:   Impact of the number of local receivers at the edge nodes packet loss rate PLR_CI

5.1.3.3 QoS/CoS guarantee and DiffServ implementation in the optical switch architecture

a. CoS/priority routing

Interestingly, there are a very simple way to support CoS and priority routing in feedback architecture. First, in the optical label, a 5-bit CoS field may be set. If an optical packet is delayed in the feedback delay line once, the CoS field will increase to a larger value according the time (i.e., traveling path length) it is delayed in the delay line. When it competes the output ports with other incoming packets, according the CoS field, the higher this value is, the higher priority it has while being forwarded.

b. Packet order maintenance

Another interesting advantage, which is very likely neglected, brought by this simple scheme shown in 1) is that the packet order is easily maintained using such scheme. As we have seen, when a packet is delayed in a feedback delay line, it will have a higher priority than other packets from the same flow. Thus, it will be possibly forwarded earlier than the later packets from the same flow. So the packet disordering in the feedback switch architecture is prevented.

c. Service differentiation

In DiffServ model, the service is provided at different classes. Traffic belonging to different class services will be granted to access different set of delayed lines to achieve different delay guarantee. Thus, service differentiation will be easily achieved.

Service differentiation in wavelength and time domains

Four classes of priorities are supported. The priority field of the packets can be set as 0, 1, 2, and 3 which corresponds best-service (BS), out-profile-assured-service (AS_O), in-profile-assured-service (AS_I), and premium service (PS) respectively. The wavelengths and delay lines resources can be accessed at different levels by the four kind of traffic. BS traffic can only access two wavelength and two fiber delay line. Both AS_O and AS_I traffic can access three wavelengths and three fiber delay lines. And PS traffic can access all four wavelengths and all four fiber delay lines. Table 1 shows the assigned resources for the four classes of traffic.

Table 4   Assigned resources

Traffic class	Priority field	Assigned wavelengths resource	Assigned delay lines
PS	3	4	4
AS_I	2	3	3
AS_O	1	3	3
BS	0	2	2

Figure 46   shows the packet loss rates for these four kind of traffic.   There's no service differentiation between AS_I traffic and AS_O traffic in these two domains, the packet loss rate of AS_I traffic is the same as that of AS_O traffic.
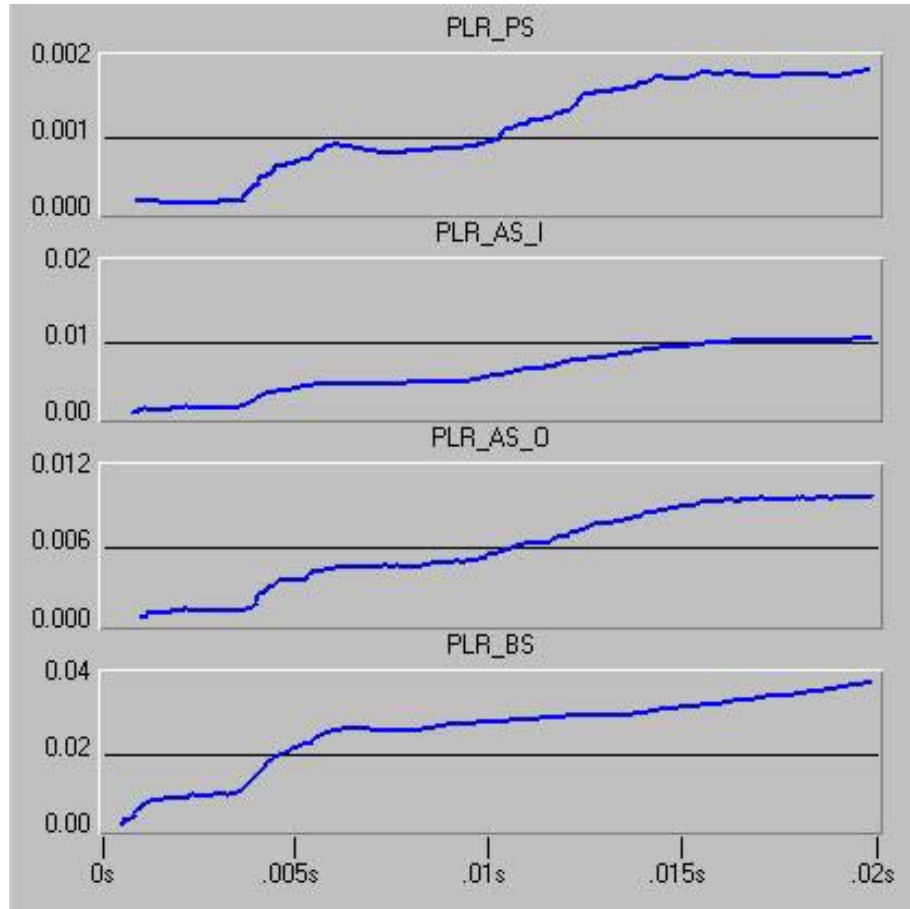
Figure 46:   Packet loss rates for four kinds of traffic

Service differentiation in electronic circulation buffers

In this section, we investigate the impact of the deployment of electronic circulation buffers as well as the RIO congestion control scheme on transmission traffic.

i). Impact of electronic circulation buffers on packet loss rate

Figure 47 and 48 show the packet loss rates with and without electronic circulation buffer.
The comparison results are summarized in Table 5.

Table 5   Packet loss rates with and without electronic circulation buffer

traffic class	packet loss rate	
	with electronic cir. Buffer	without electronic cir. buffer
Premium Serv. Traffic (PS)	0.040%	0.052%
Assured Serv. Traffic (AS)	0.38%	0.43%
Best Effort Serv. Traff. (BS)	1.05%	1.25%
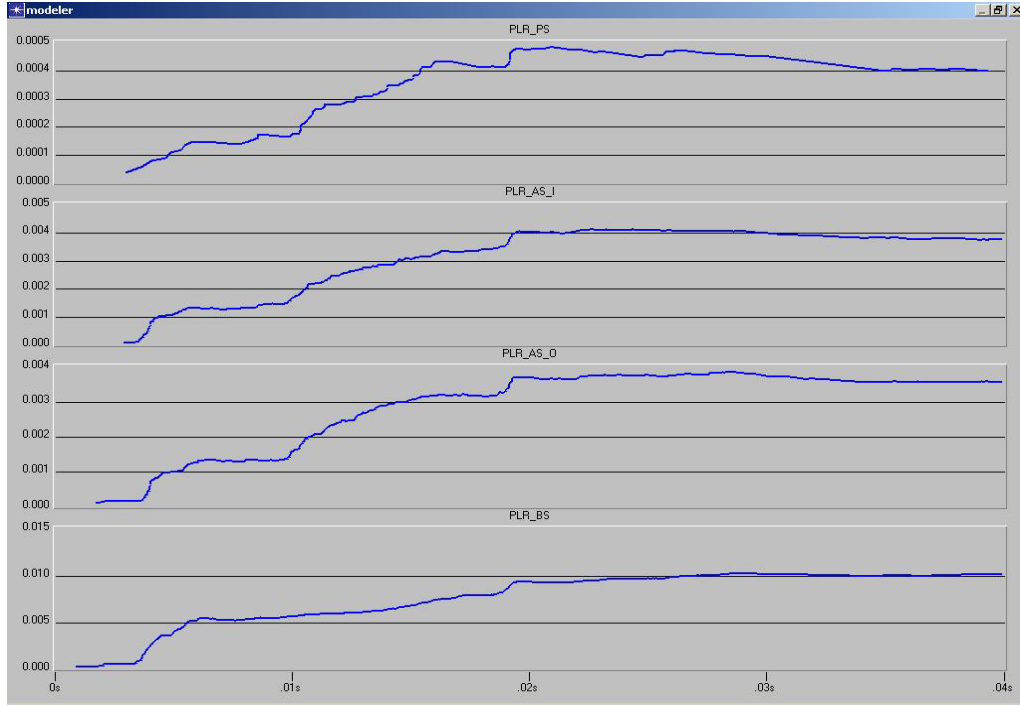
Figure 47: Packet loss rates with electronic circulation buffer
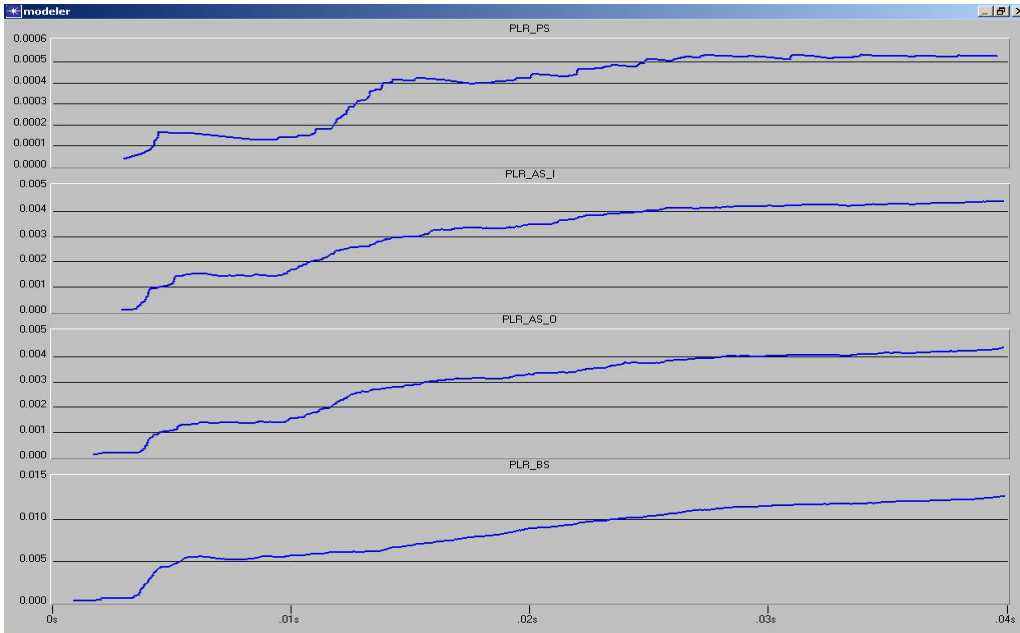
Figure 48: Packet loss rates without electronic circulation buffer

ii). Impact of RIO congestion control scheme on packet loss rate

In our simulation, the traffic generator generates 30% PS traffic, 20% out-profile AS traffic, 20% in-profile AS traffic, and 30% BS traffic. Three kinds of queues are used to manage the traffic: PS queue, AS queue, and BS queue. RIO algorithm is used to manage AS queue. FIFO schemes are used to manage PS queue and BS queue. A static priority queuing scheme is used to manage these three queues. The

packets with higher priority are served earlier than lower priority traffic according static priority scheduling scheme. Figure 49 shows the RIO algorithm. It is a variant of RED algorithm.
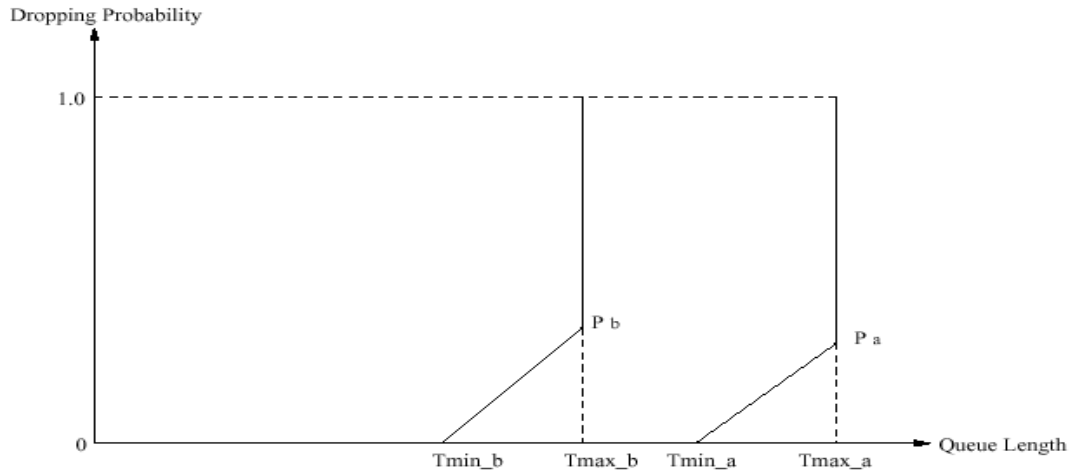
Figure 49:   RIO algorithm
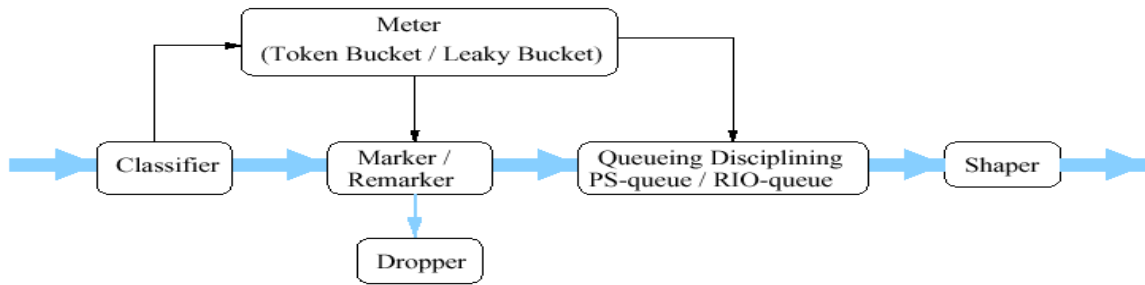
Figure 50:   Function models in electronic circulation buffer management unit

Figure 50 shows the functional models for service differentiation in the electronic circulation buffer. Figure 51 shows the buffer management schemes used in our simulation. In the future work, CBQ or WFQ can be used instead of PQ for better system performance further.
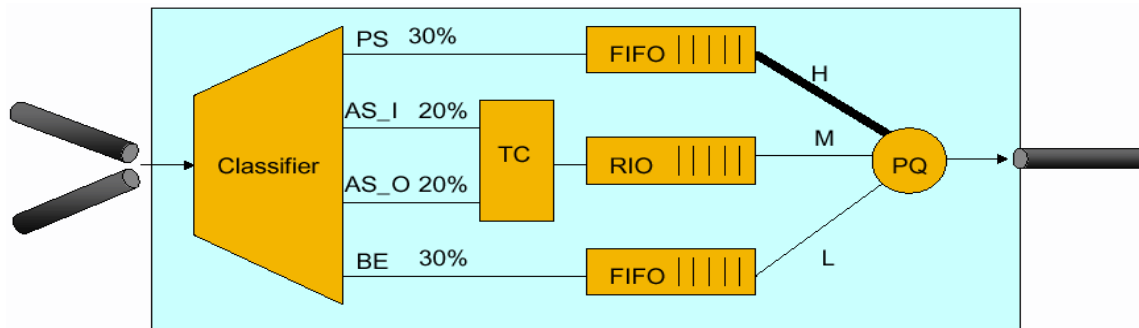
Figure 51:   Buffer management schemes

Figure 52 shows the packet loss rates for these four kinds of traffic.   Figure 53   shows the impact of the RIO algorithm on out-profile traffic and in-profile traffic. Here a shorter length AS queue is used to shorten the real simulation time.
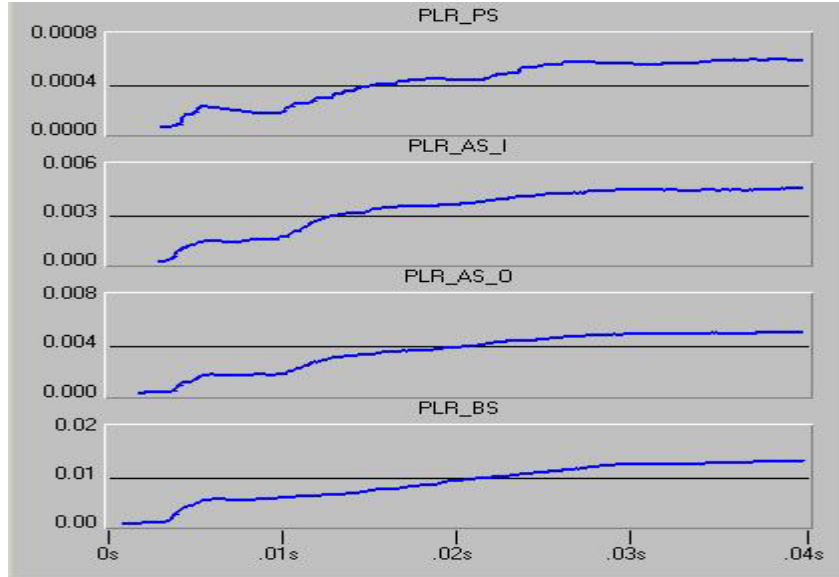
Figure 52: Packet loss rates for different classes of traffic
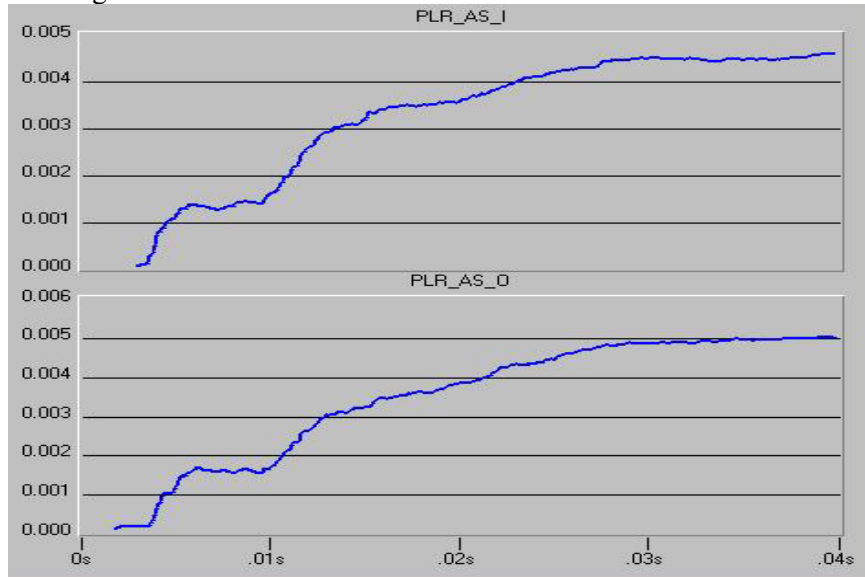
Figure 53: Impact of RIO on AS traffic

5.2 Optical-Label Switching Router Table Integration
5.2.1 Client Interface

5.2.1.1 Experiments

Client interface experiments for both one hop and two hops switching with label swapping were done. The client interface acted as a one input port, one output port edge router. An optical-label switching (OLS) edge router's main function is to seamlessly interface networks that use different protocols by generating or extracting optical labels to be used in the OLS core network. Figure 54 shows an example used in this experiment for interfacing packet over SONET (POS) client networks and an OLS network via edge routers. The figure shows the edge router functionality along with the ingress and egress

paths to/from the OLS network. When the edge router receives packets on its POS interface, or ingress path, the physical layer interface converts the optical signal to an electrical signal. The edge router terminates the SONET frames leaving PPP-encapsulated IP packets. The packets are sent to the data bus traffic controller in preparation for forwarding across the data bus to the edge router's OLS interface. The OLS interface's data bus traffic controller receives packets from the data bus and passes them to the label-processing unit. The label-processing unit uses the contents of the IP header to generate a label that is used by the OLS core routers to make the forwarding decisions. The label contains a preamble, egress edge router destination address derived from the destination IP address, priority derived from the IP type of service field, packet duration derived from the IP datagram length, optical time-to-live (OTTL) derived from the IP TTL field, and the source address of the ingress edge router. The label at 155Mbps and the PPP-encapsulated IP packets at 2.488Gbps are then forwarded to the physical layer interface for sub-carrier multiplexing (SCM) and transmission to the OLS network.
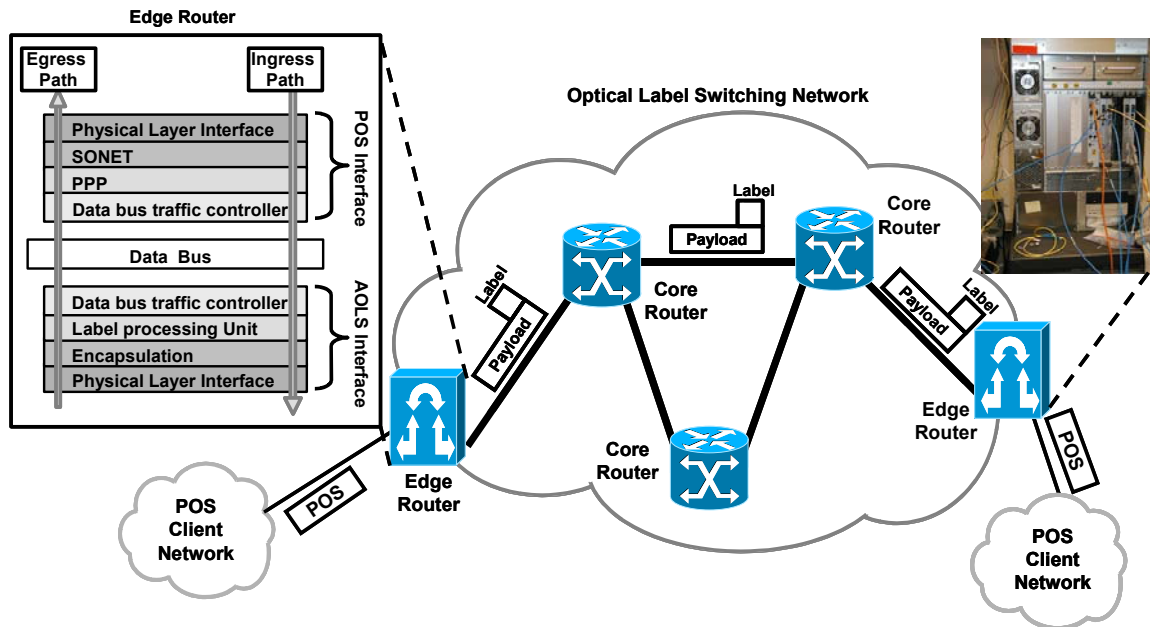
Figure 54: Overall network highlighting edge router functionality. The ingress path starts at the POS network, enters the edge router's POS interface, the OLS interface and then the OLS network. The egress path starts in the OLS network, enters the edge router's OLS interface, the POS interface and then the POS network.

When an OLS core router receives an optical packet, it uses all-optical label extraction and achieves decoupling of the control and the data planes. The router controller uses the contents of the extracted labels to perform the forwarding decision and forwards the packets towards their egress edge router. Since optical queuing buffers are not available, the OLS core router performs all-optical packet contention resolution using a wavelength-, time-, and space-domain switching scheme. At the egress edge router, the OLS interface's physical layer interface receives the encapsulated packets. Since the OLS network is asynchronous, the incoming packets must be byte-aligned. The byte-aligned packets are sent to the data bus traffic controller for forwarding across the data bus to the POS interface's data bus traffic controller. The POS interface encapsulates the IP packets in PPP, frames them in SONET, and forwards them to the POS network. Currently, the edge router has one input port and one output port. The majority of the edge router design was implemented in field programmable gate arrays (FPGAs) allowing for design flexibility and short development time.

Figure 55 shows the experimental setup for the two-hop IP client-to-IP client communication using edge routers and OLS core routers. The IXIA OC-48c POS Load Module emulates the POS client

network. It generates and sends three different 1500 byte IP packets, P1, P2, and P3 to the ingress edge router. Each kind of packet has a different destination IP address. The edge router reads the IP headers, generates labels L1, L2, and L3 at 155Mbps, and places them on the three packet payloads, P1, P2, and P3 at 2.488 Gbps, by sub-carrier multiplexing (SCM), thus creating three kinds of optical packets. The OLS core router employs all-optical label extraction in label extractor 1 and achieves separation of the control and the data planes. The first burst-mode receiver (BM Rx1) recovers the contents of labels L1, L2 and L3 and sends them to the FPGA based router controller, while the payloads P1, P2, and P3 are amplified and delayed using an erbium-doped fiber amplifier (EDFA) to match the ~260 nsec processing time through the router controller. Using the recovered labels, the router controller decides where to forward the packets and sends the appropriate control signals to the first tunable laser diode (TLD1). TLD1 with the semiconductor optical amplifier (SOA2) converts P1 and P2 to 1552nm and P3 to 1546nm as dictated by the router controller based on the contents of L1, L2, and L3. The two wavelengths then assume the switching paths determined by the well-known wavelength routing characteristics of the arrayed-waveguide grating router (AWGR). This results in dropping P3 at first hop drop port and forwarding P1 and P2 to the EDFA1 where gain clamping is used to reduce the gain transient. The router controller generates two new labels, L1' and L2', updating the contents if necessary. In the label-swapping module , P1 and P2 are sub-carrier multiplexed with their new labels L1' and L2'.

The optical packets with new labels now enter the second hop through label extractor 2 where the OLS core router performs forwarding similar to the first hop with a Mach-Zehnder interferometer wavelength converter (MZI WC) used in place of the SOA. The router forwards P2 to a drop port (1542nm) and forwards P1 to the receiver and egress edge router (1546nm). The egress edge router receives P1, performs byte alignment, converts the received OLS packet to SONET frames and sends them back to the IXIA OC-48c POS Load Module for performance analysis.
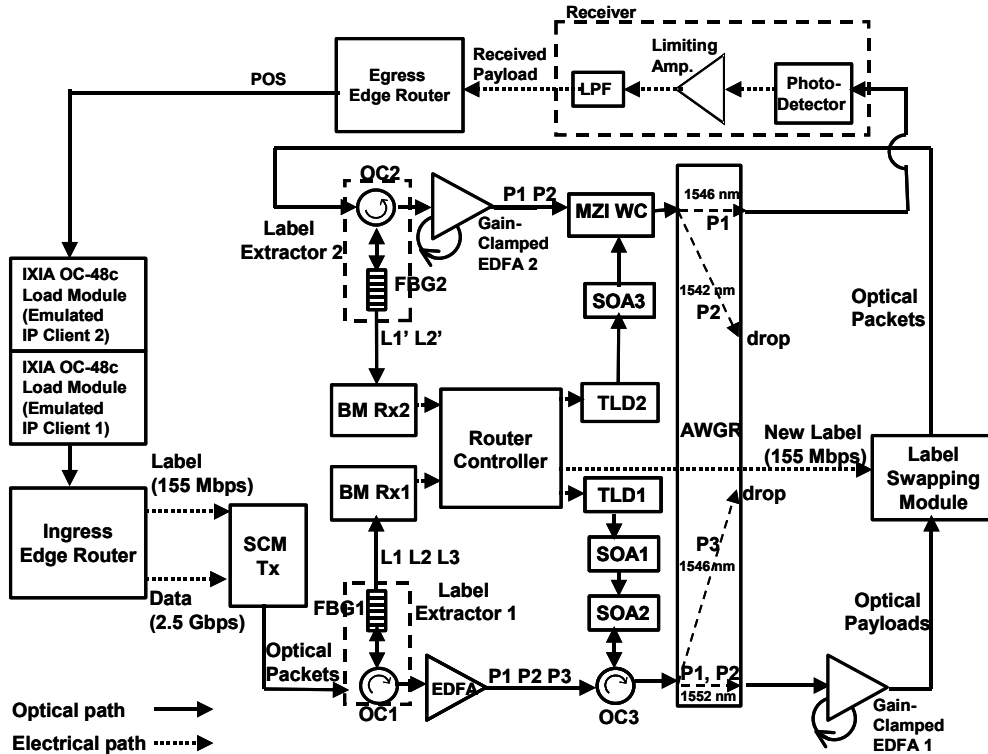
Figure 55:   Experimental setup for two-hop IP client-to-IP client communication using edge routers and an OLS core router.

Figure 56 shows the measured packet-error-rate (PER) and calculated bit-error-rate (BER) along with packet patterns and eye diagrams. The experimental setup with IP-over-optical using optical label switching does not offer layer 1 or layer 2 testing since the edge router functions as a layer 3 device resulting in no layer 1 or 2 testing access. Hence, in this experiment, an estimation of the BER based on the measured PER is used. The estimation utilizes the formula

$$PER = 1 - (1 - BER)^N,$$

where N is the number of bits in a packet. This estimation assumes uniform error rates for all bits, regardless of their position within the packet, and the probability of an error is the same for all bits in a memoryless system. The crossing in the PER and BER curves are due to reduction in average optical power after dropping P3, or P2 packets. The results show a low power penalty at BER = 1E-9 for two-hop switching. These results also show that the edge router - OLS core router combination works well together with a good performance.
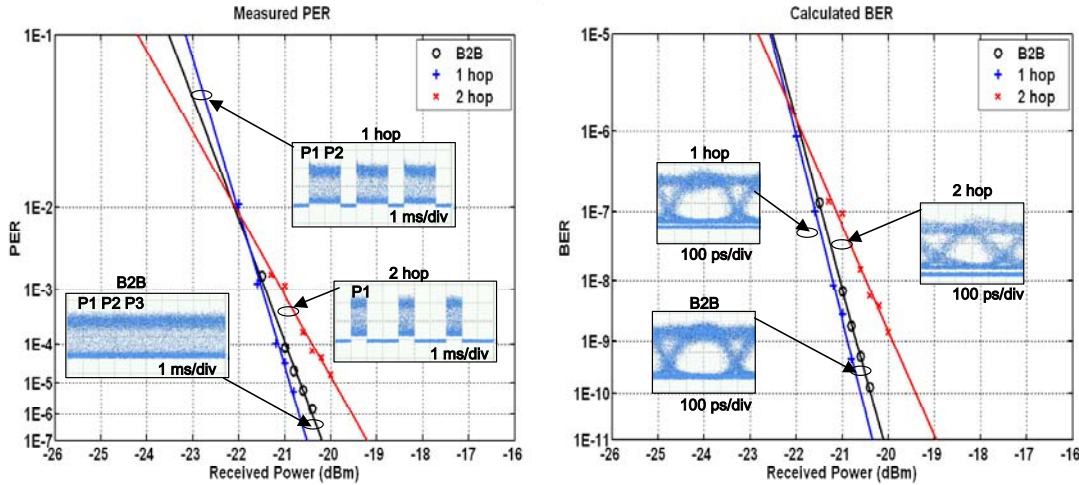
Figure 56:   Experimental results for two-hop IP-client to IP-client communication using edge routers and an OLS core router (a) Measured packet-error-rate (PER) with packet patterns (b) Calculated bit-error-rate (BER) with eye diagrams.

We demonstrate, for the first time, the transport and routing of IP traffic from IP client-to-IP client over a two-hop OLS core router. Edge routers interface the IP clients and the OLS router with low packet-error-rates.

5.2.1.2 Client Interface Algorithms

Figure 57 shows the block diagram of the client interface and it's connection to the client network via the edge router. The system has five main stages: edge router, input queues, request polling, scheduler, and output queues.

50
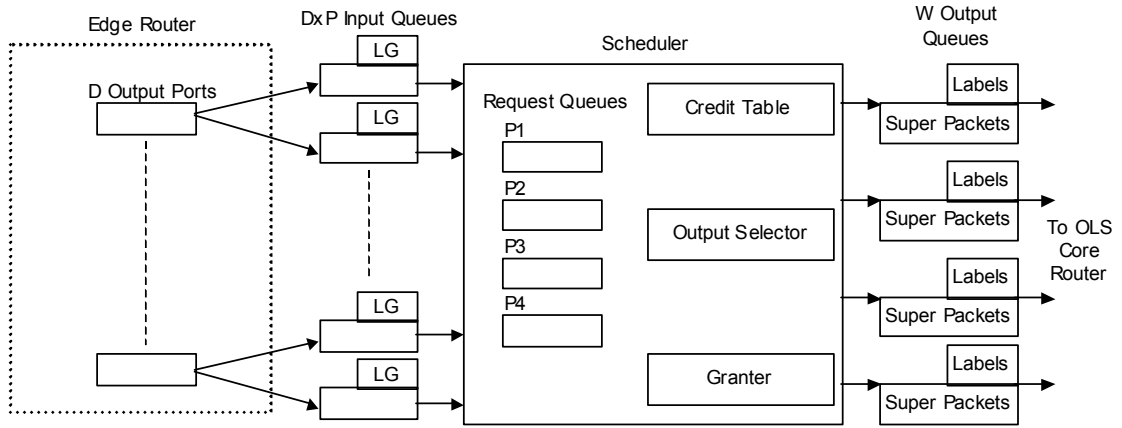
Figure 57:   Block Diagram of Client Interface with Edge Router Connection

a. Edge Router

The edge router is responsible for sorting the packets destined for the OLS network into D separate output ports. The edge router's output ports correspond to the D possible destinations in the OLS network. NC&M will maintain the forwarding table in the edge router that maps the client packet IP address to the OLS network destination. The edge router outputs the destination-sorted packets to the client interface.

b. Input Queues

The client interface sorts each port's packets into P queues according to their priorities. Therefore, at the client interface input there is DxP input queues as shown in Figure 57.   These queues will be used for the packet aggregation algorithm presented in [1]. The queues accumulate IP packets to make super packets.

The active queue management algorithm GREEN will be adapted to manage each of the DxP queues. GREEN manages the queues based on the estimated input arrival rate and not the queue length as RED does. GREEN maintains significantly shorter queues than RED [2] and thus it is believed that GREEN will be more suitable for use with the packet aggregation used in the client interface.

Since packet aggregation attempts to grow the queue and GREEN attempts to shorten the queue, the two algorithms are working against each other. Therefore, the client interface may perform better if the two algorithms are combined to operate together.

Combined Algorithm

The flow chart in Figure 58 illustrates the packet arrival process with the combined packet aggregation and modified GREEN algorithm at each input port. The modified GREEN algorithm is called rate estimation (RE) and is explained below.
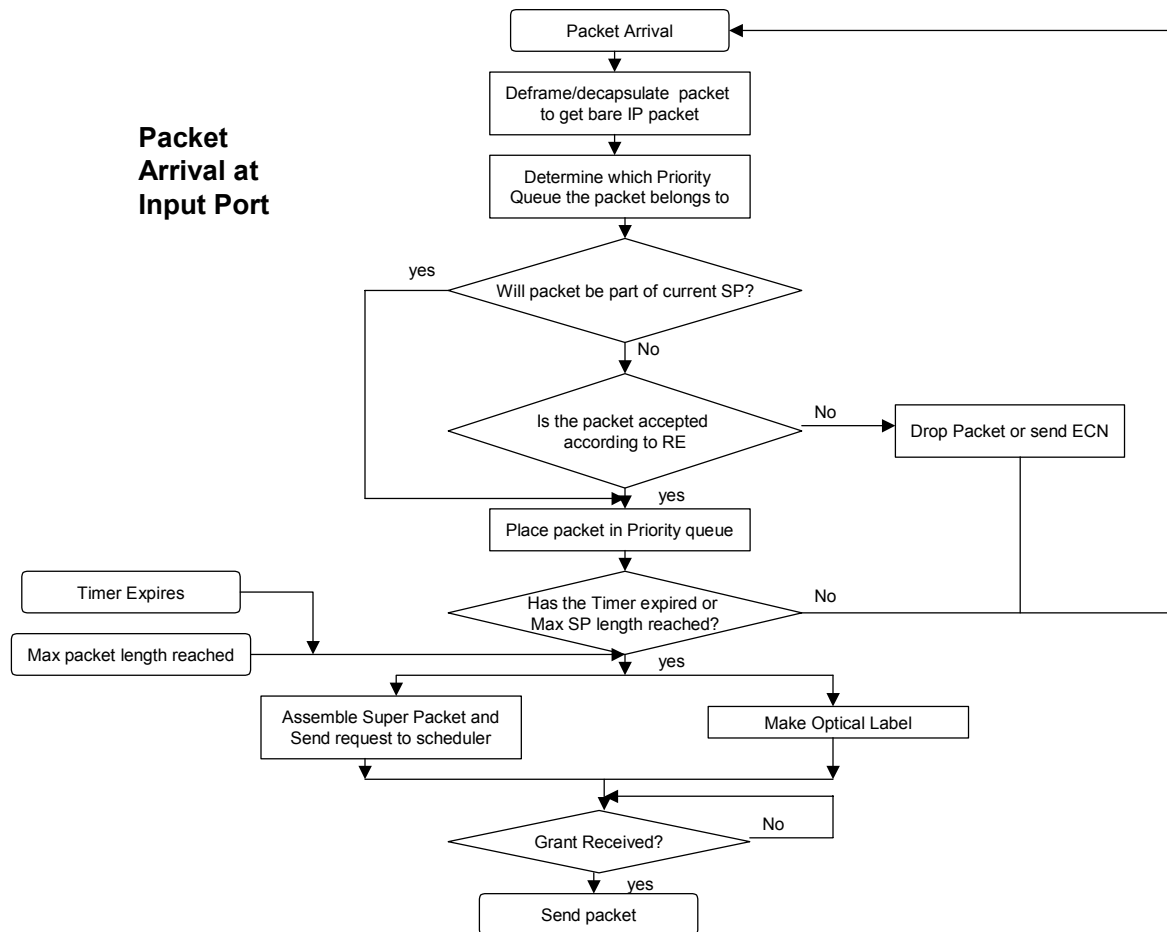
Figure 58:   Flow chart of the combined packet aggregation / RE algorithm at the input queues.

First, the packet is deframed or decapsulated to get bare IP packets. The priority of the incoming packet is then identified to determine which queue the packet belongs to. If the incoming packet will be part of the current super packet, the packet is accepted. This is done so that the packet aggregation algorithm can operate with a minimum length of one super packet without the RE algorithm limiting it.   If the packet will not be part of the current super packet, the packet is accepted according to the RE algorithm. If the RE algorithm rejects the packet, the packet is dropped or explicit congestion notification (ECN) is sent to the sender. If the packet is accepted according to RE, the timer and maximum super packet length are checked to determine if the super packet is ready to send. If it is time to send the super packet, a request is send to the scheduler and an optical label is assembled. Once a grant is received from the scheduler, the super packet and label are sent to the arbiter and subsequently the output port.

It is expected that the input priority queue lengths will grow beyond the super packet size since the packet arrival rate may exceed the scheduler's grant rate. Therefore, the incoming packets must be queued. This may lead to a situation where enough packets to make multiple super packets are in the input priority queue.   A process watches the queue to determine if this happens.   This is shown in Figure 58 by the max packet length reached process block. Figure 55 also shows the timer expired process block.   This process monitors the timer and if it expires, the super packet is assembled and a request sent to the scheduler.

Rate Estimation (RE) Algorithm

The rate estimation algorithm is similar to the GREEN algorithm presented in [2] and adapts part of the algorithm presented in [3]. It adjusts the rate of congestion notification by using an estimate of the incoming data arrival rate, x_i, as a feedback control function. If the estimated arrival rate is above the target capacity, c_i, the rate of congestion notification, P, is incremented by ΔP at a rate of $1/\Delta T$. If the estimated rate is below the target capacity, then P is decremented by ΔP at a rate of $1/\Delta T$. Probabilistic marking of the incoming packets at a rate P is applied. P is calculated as follows:

$$P = P + \Delta P * U(x_i - c_i)$$

Where

$$U(y) = \begin{cases} +1 & y \leq 0 \\ -1 & y < 0 \end{cases}$$

The rate estimation algorithm is applied on each of the DxP input queues independently. Therefore, the estimated arrival rate for input queue i is x_i, and its target capacity is c_i. The estimated arrival rate is calcaluated as follows:

$$x_{inew} = (1 - e^{-Del_i/K}) * \frac{l_i}{Del_i} + e^{-Del_i/K} * x_{iold}$$

Where K is a constant, Del_i is the inter-packet delay, and l_i is the packet size for input queue i. The constant K should be between 100 and 500ms [3]. A large K value better filters the noise and avoids potential system instabilities. A smaller K value increases the system responsiveness to rate fluctuations. More on the selection of K can be found in [3].

The target capacity for each input queue, c_i, can be different for each queue. The queues with high priority can have a higher target capacity than those with low priority. The target capacities are sent by NC&M using the knowledge of the entire network. ΔT, the rate as which P is updated, needs to be large enough to allow changes to take effect before more changes are made. Thus, ΔT should be larger than the average round-trip-time (RTT) of the network. Estimating the average RTT can be very difficult to do dynamically; therefore, a static value will be used. According to derivations in [4], 10ms is a good average RTT.

Label Generation

The label is generated at the input queues since most of the label contents are known at the input queues. In Figure 57, the LG units at each input queue illustrate the label generation. The source, destination, priority, and ttl are all constant for each input queue. The packet length is the only computation that is needed.

c. Request Polling Arbitration

Once a super packet is ready to be sent, the input queue sends a request to the appropriate priority request queue in the scheduler. Since there are DxP input queues and only P priority request queues, arbitration is needed for entry into the priority request queues. The arbitration will be done in a manner similar to [5].

d. Scheduler

Because all the outputs go to the same OLS core router, the scheduler is not a traditional scheduler. Any request sent to the scheduler can be sent to any of the output queues. Therefore, one of the scheduler's main goals is to achieve load balancing between the W output wavelengths. Figure 57 shows the basic functional blocks of the scheduler. It has P polling queues representing the P priorities.   The scheduler decides which queue to service according to a credit table. The credit table tells the scheduler which queue to service and how often to do so. The scheduler will continuously monitor the output queues to determine which output the packet is sent to. This can be done in two ways:

1. round robin
2. shortest length with round robin tie breaking

Simulations can be done to determine which one provides the best load balancing. After the output queue is selected, the scheduler grants the request and places the packet and label in the output queue.

e. Output Queues

In Figure 57, W output queues are shown for the client interface.   The W outputs queues correspond to the number of wavelengths connecting the client interface and the OLS core router. All the output queues go to the same OLS core router; the only difference is the wavelength. The task of the output queues is to output super packets and their label to the OLS network.

References

[1] Fei Xue, Shun Yao, Biswanath Mukherjee, and S. J. B. Yoo, "The performance improvement in optical packet-switched networks by traffic shaping of self-similar traffic," Technical Digest, IEEE/OSA Optical Fiber Communication Conference, paper #WG4, Anaheim, California, (2002).

[2] Bartek Wydrowski and Moshe Zukerman, "GREEN: An Active Queue Management Algorithm for a Self Managed Internet", Proceedings of ICC 2002, New York, vol. 4, pp. 2368-2372, 2002.

[3] I. Stoica, S. Shenker, H. Zhang, "Core-stateless fair queueing: achieving approximately fair bandwidth allocations in high-speed networks," Computer Communication Review, vol.28, (no.4), ACM, Oct. 1998. p.118-30.

[4] F. Callegati, and C. Raffaelli, "End-to-end delay evaluation for an optical transparent packet network," Photonic Network Communications, vol.1, (no.2), Kluwer Academic Publishers, Aug. 1999. p.147-60.

[5] Yash Bansal, "Asynchronous, unslotted optical packet switching router controller with priority based traffic switching," M.S. Thesis, Sept. 2002 .

5.2.2 Performance Monitoring for optical label switching network.

5.2.2.1 Background

 Performance monitoring is a important aspect in optical communication systems. If effective performance monitoring and protection and restoration can be provided in optical layer, it will facilitate the progress of the evolution of communication network toward the architecture of IP over WDM or multi-protocol over WDM. Currently optical network performance monitoring is based on some indirect measurements such as histogram of the optical signal, optical power, extinction ratio, signal to noise ratio, and so on. The basic requirement for optical performance monitoring is that it is not intrusive to the signal. And it is preferred if no OE-EO conversion is necessary for the signal. Currently several methods are proposed for optical network performance monitoring:

- Measure several aspects of signal qualities and estimate BER. In this method BER is derived indirectly from optical signal quality indicated in power, SNR, extinction ration, and in some cases dispersion
- Synchronously sample optical signal and build histograph of the optical signal then deduce BER by extrapolation.
- Asynchronously sample optical signal and estimate the BER based on histograph.

This research work proposes use optical label as an indication of payload signal quality and monitoring payload BER by monitoring label BER in sub-carrier channel. The performance monitoring experiment is based on the assumption that the label BER and the payload BER in an optical label switching network where labels are carried in sub-carrier. The experiment investigated relationship between the label BER and the payload BER at different signal to noise ratios.

The experiment set-up is shown in Figure 59 the payload and label signal is combined in the microwave combiner and then modulate the optical signal from a DFB laser. Then the signal is sent through EDFA. By adjusting the attenuator before the EDFA, signal to noise ratio can be controlled. After the EDFA, BER is measured for payload and label simultaneously. Payload is extracted by a FBG grating.

PM Experiment Set-up
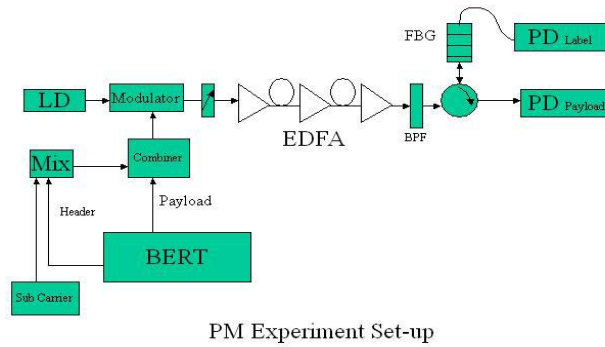
Figure 59: PM Experimental Set up

The payload and label BER are shown in Figure 60, label BER is denoted by L and payload BER is denoted by P. Their correlation is shown in Figure 61. Each curve in this graph corresponds to an attenuation level before the EDFA, so each curve represents a SNR level of the optical signal. The experiment results show that the BER of payload data is closely related to that of label. But analysis of the data shows that the relationship between these two BERs is not a simple one. The correlation is coupled with signal to noise ratio. So that it is not possible to accurately decide payload data BER only according to packet label BER in the sub-carrier channel. The signal to noise ratio must also be known to estimate BER of payload. But the experiment only took into account the influence of signal to noise ratio, other factors such as dispersion, nonlinearity are not taken into account. So that the correlation inherently is not a simple relation and it is not possible to accurately estimate the payload BER merely according to label BER.
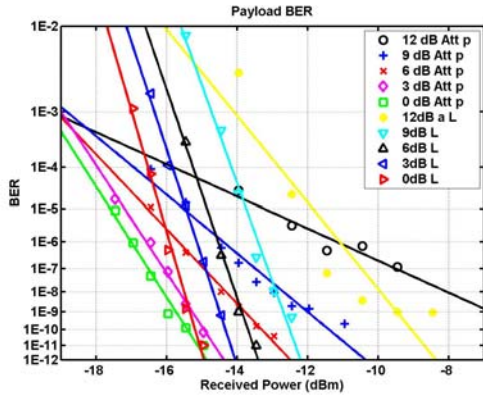
Figure 60: payload and label BER	Figure 61: correlation between payload and label
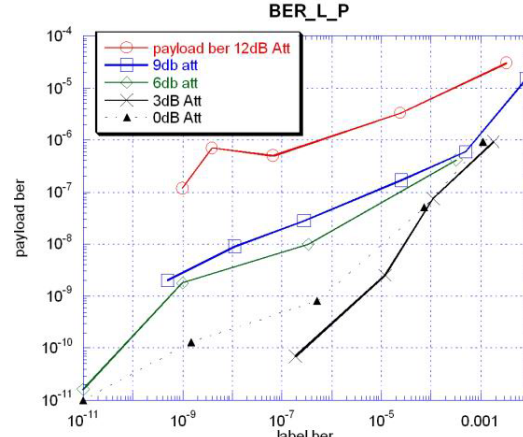
In summary, the experiment showed correlation between the payload and label BER, but the correlation is also coupled with the impair factors of the signal. So the impair factor measurement such as SNR is also need in estimation of payload BER.

New method of optical performance monitoring is needed to make it practically useful. One of the possible solutions is to add some overhead traffic over the optical label switching network, and use the overhead to measure transmission quality.

5.2.3. Optical core router controller

In order to estimate the performance in the controller architecture, an analytical model based on the hardware implementation is developed. As with any architecture, the performance is measured in latency in the controller. Additional benefits of this model include – scalability estimations and performance estimations with higher device technology. The development of the analytical model is also presented.

5.2.3.1 Architecture support of label-rewriting

Previous optical packet routing system demonstrations had been for a single hop system. Based on experimental observations it was clear that both the current cross-gain and cross-phase modulation based wavelength conversion schemes caused the sub-carrier label signal to deteriorate. Thus, after wavelength conversion, the burst mode receivers could not detect the label. For a network of optical packet switching routers and multi-hop packet forwarding, it is essential to read the label at each router. The solution was to include label rewriter module at each output of the switch fabric. During the forwarding process, each label field is registered inside the controller. As a result, each label-rewriting module receives an electrically regenerated label input from the controller. This also allows update of several fields such as optical time-to-live (OTTL), and a modified label can be sent to the label rewriter module.

For forwarding each packet, the controller receives and detects a new label, performs forwarding table look-up, resolves any packet contention, sends control signals to tunable wavelength converter (T_WC) and additionally after some delay sends electrically regenerated label to the label-rewriter module at the output of the switch fabric. Figure 62 shows the controller receiving labels from the BM_RX, routing control signal being sent to the T_WC and regenerated label from the controller being sent to label-rewriting module (L_WR). With the label-rewriting module, once the forwarding decision is made, the controller first sends control signal to the T_WC to route a packet and after a fixed delay δ (in clock

cycles) sends a serial label signal to the L_WR. The fixed delay δ is added to compensate the propagation delay of the payload signal from T_WC to L_WR.
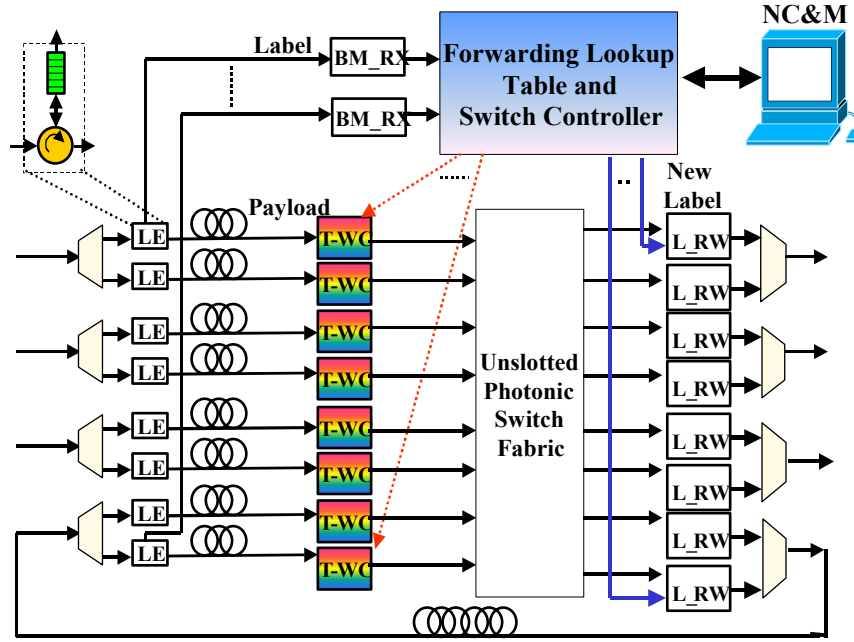
Figure 62:   Optical router system diagram. LE decouples the control plane from the data plane. Controller receives label from BM_RX, performs forwarding table lookup, sends control signal to T_WC to route packets through switch fabric. New label from controller is fed to L_WR.

The controller receives label at 155MHz and stores the various label fields in appropriate registers. Once the label fields are registered, the forwarding process is performed using a 77.5MHz clock. For the label rewriter stage, the label needs to be serially transmitted at 155MHz. The forwarding process for a packet is complete after the control signal is sent to T_WC. The T_WC also requires an enable signal to indicate a valid control signal. This enable signal Enb, is also used to trigger the label transmit process inside the controller. Figure 63 shows the state-machine, triggering on enable and reading the 2-bits from the label registers in each state. The state-machine waits for δ clock cycles in state S1 after receiving the enable signal. Between states S2 and S14, the label bits are read 2-bits at the time from the label registers. The preamble field is read in states S2-S5, followed by the destination fields in states S6-S7. The priority field is read in state S8, with the packet length field read in state S9-S10. States S11-S12 read OTTL field and finally in states S13-S14, the source field is read. This state-machine operates at 77.5MHz, while the label need to be transmitted serially at 155MHz. Thus, the 2-bit output from each state is fed into a 2-input multiplexer that changes its select lines at 155MHz. The resultant output of the multiplexer is a serial label at 155MHz.

These additions to the controller architecture complete the design of the control plane for an optical packet switching system.

Reset = 1

S0 **Enb = 1** **δ > 0**

S14 O_{78}=Src[1:0]

S1 **δ = δ -1**

S13 O_{78}=Src[3:2]

δ = 0

S2 O_{78}=Pre[7:6]

S12 O_{78}=TTL[1:0]

S3 O_{78}=Pre[5:4]

S11 O_{78}=TTL[3:2]

S4 O_{78}=Pre[3:2]

S10 O_{78}=Plen[1:0]

S5 O_{78}=Pre[1:0]

S9 **S8** **S7** **S6** O_{78}=dest[3:2]

O_{78}=Plen[3:2]

O_{78}=Pri[1:0]

O_{78}=dest[1:0]
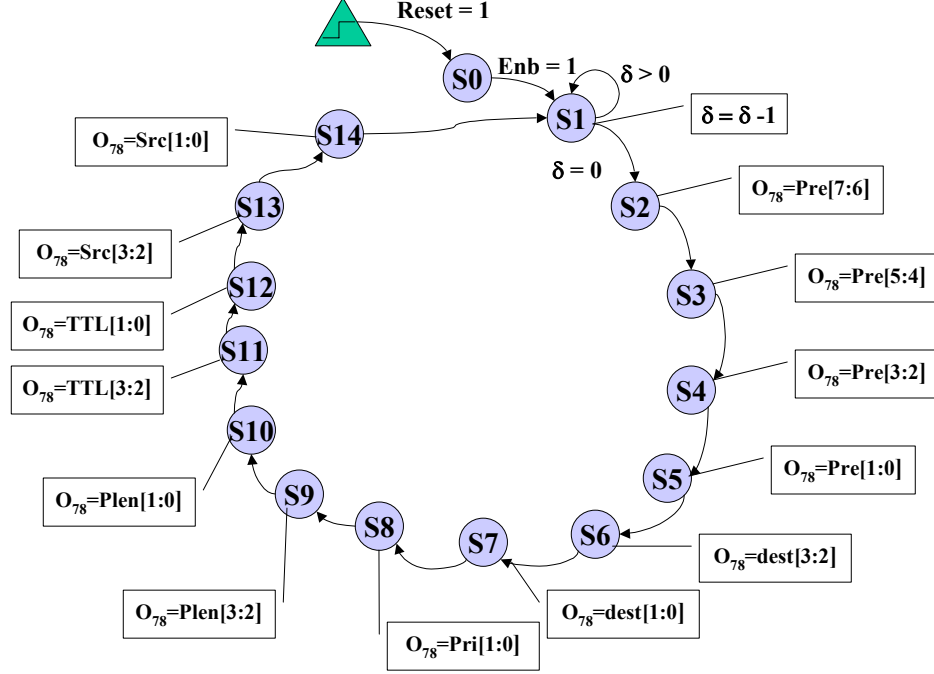
Figure 63:   State-machine running at 77.5MHz, and transmitting the label fields 2-bits in each state

5.2.3.2. Analytical model for control plane

The analytical model currently assumes a controller supporting 4 fiber inputs with W wavelengths/fiber. Two fibers are used as transport fibers connecting other optical routers, 1 fiber for fiber delay line and 1 fiber for the local client add/drop traffic. However the model can be extended for any number of fiber inputs and wavelengths. The figure of merit for this model is average forwarding latency, i.e., the average number of clock cycles required to complete the forwarding of a packet. The average forwarding latency $\Delta_{latency}$, through the router controller is the summation of the latency of through the three main components of the architecture – *IChannel*, arbiter and *scoreboard*.

$$\Delta_{\text{scoreboard}} = \Delta_{\text{IChannel}} + \Delta_{\text{arbitrate}} + \Delta_{\text{scoreboard}} \qquad \ldots(1)$$

Figure 64   shows the block diagram of the control-plane and indicates the three main components, *IChannel*, *arbiter* and *scoreboard*, in the controller architecture. These three components have been described in detail in previous progress reports. This report will present the analytical modeling of these three sections.

The first section in the architecture is the *IChannel*. The function performed by each *IChannel* is analogous to that performed by a line card in an electronic router. It performs label field processing, forwarding table look-up and make requests to be routed on the preferred output port through the arbiters. However unlike line cards in conventional routers, the *IChannel* uses the label to perform the forwarding decision and the label bit-rate is significantly lower than the payload bit-rate. The asynchronous, unslotted switching of packets also avoids the need for complex segmentation and reassembly of packets. As a result *IChannel* can be built with simple electronics inside the controller.

58
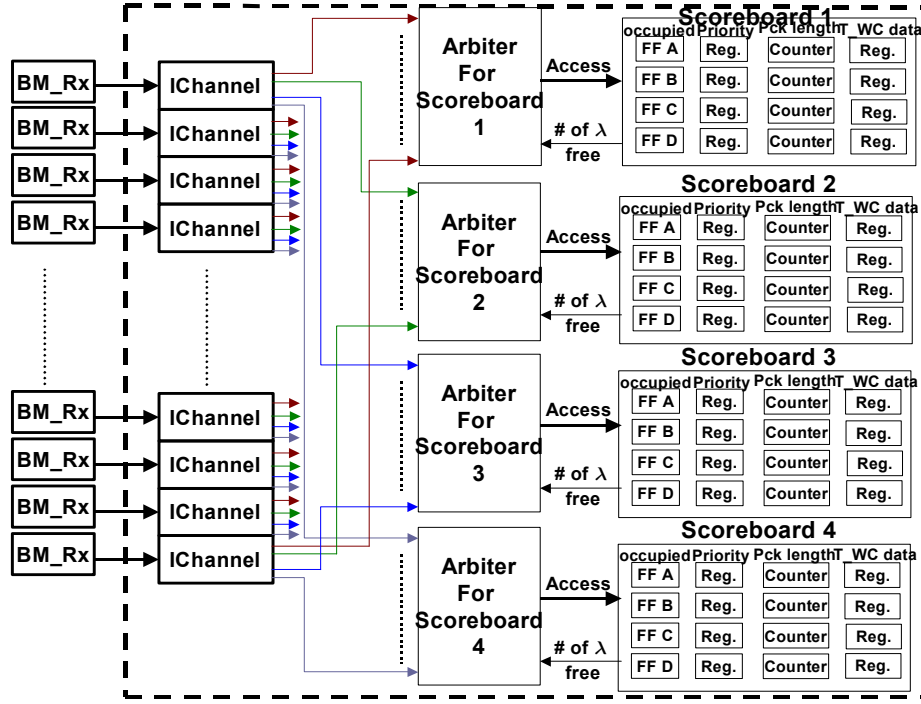
Figure 64: OLS router Controller internal architecture

The *IChannel* does not wait to receive all the label fields, and starts making forwarding requests to the output port arbiters as soon as the RAM look up is complete. Therefore, the clock cycle latency through the *IChannel* is fixed and depends upon preamble detection, destination lookup on the routing table for the first and second preferred paths, and storing the results of the routing table lookup.

$$\Delta_{\text{IChannel}} = \frac{(P+D)}{LR} \times CR + Ram\_lookup + register \qquad \dots(2)$$

Where, P = Number of Preamble bits, D = Number of Destination Bits, LR = Label Rate, CR= Clock Rate

The next component in the controller architecture is the arbiter. The calculation of the arbitration latency is complex because the latency is dependent upon the traffic in the system, i.e., number of wavelengths occupied by other traffic on the each path. Further, the probability of being accepted/denied on each path based on the contention resolution needs to be calculated. The result is a variable arbitration latency depending upon which path a packet is accepted and forwarded on. Each arbiter has identical operation, and accepts 1 requestor to be routed on the corresponding output port, during each clock cycle, until all *W* wavelengths on the path are occupied. Let P_a be the probability that an arbiter accepts a request. Once all wavelengths are occupied, the arbiter denies all subsequent requests till a wavelength becomes free. Let P_b the probability that all wavelengths are occupied, and an arbiter denies the request. If a request is denied after *W* attempts, the request is lost, as there is no queuing of requests and a similar request is made to the next arbiter in the contention resolution algorithm. The process of acceptance and denial on

each arbiter is shown in Figure 65. Let the arrival rate R packets/sec, for every destination, be identical on each wavelength/fiber and follow the Poisson process. Then $\lambda = (N \times W \times R)$ packets/sec be the rate of arrivals on all NxW inputs. The packet size has a negative exponential distribution of S-bits, thus duration of each packet is $1/\mu = S/DR$ sec, where DR= payload data rate in Gbps.

(a) **(b)**
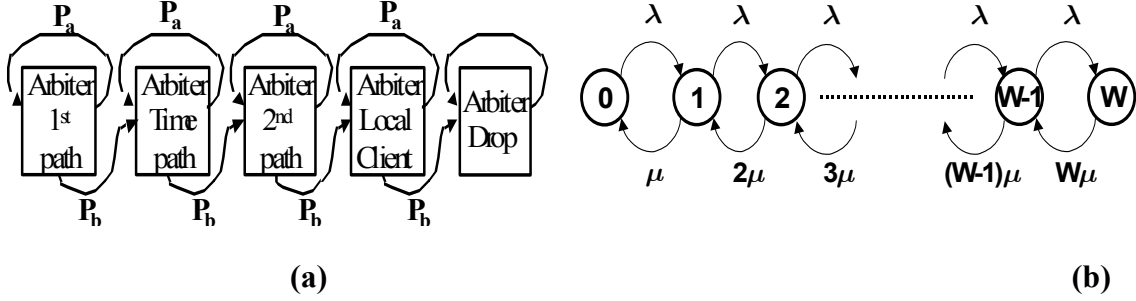
Figure 65: The acceptance and denial through various arbiters (a). Discrete time Markov chain for M/M/m/m system (b). State W corresponds to W customers in the system

Each arbiter can be considered equivalent to an *M/M/m/m* system (Figure 65 (b)), with *m* servers and at most *m* requestors currently in service (where *m* = *W* is the number of wavelengths in the fiber). This model does not allow queuing of requests. Equivalently, this model can be used in the OLS router for each arbiter, which receives requests for wavelengths on an output fiber and the maximum number of wavelengths allowed is *W*. The probability of being accepted on an arbiter is equal to the probability that there is at least one free wavelength on the output fiber and the arbiter system is between state 0 or state W-1.

$$P_a = P_0 + P_1 + P_2 + \cdots\cdots + P_{W-1}$$

$$P_a = \left[\sum_{k=0}^{W-1} \left(\frac{\lambda}{\mu} \right)^k \frac{1}{k!} \right] P_0 = \frac{\left[\sum_{k=0}^{W-1} \left(\frac{\lambda}{\mu} \right)^k \frac{1}{k!} \right]}{\left[\sum_{k=0}^{W} \left(\frac{\lambda}{\mu} \right)^k \frac{1}{k!} \right]} \qquad \ldots(3)$$

Since each arbiter accepts only 1 request during each clock cycle, in the worst case each requestor sends *W* requests before being sure of denial. The latency of acceptance depends upon in which state the request is accepted.

$$\textit{Latency of acceptance on arbiter} = L_{accp} = P_k(k+1) = \frac{\left[\sum_{k=0}^{W-1} \left(\frac{\lambda}{\mu} \right)^k \frac{1}{k!}(k+1) \right]}{\left[\sum_{k=0}^{W} \left(\frac{\lambda}{\mu} \right)^k \frac{1}{k!} \right]} \qquad \ldots(4)$$

Since, the M/M/m/m system does not allow queuing of requests, at any instant of time the system has to be in one of the states of Figure 65b. The probability P_b that a request is denied by the arbiter equals the probability that all W wavelengths are occupied on the output port. This can be calculated by the Erlang B formula.

$$P_b = P_w = (1 - P_a) = \left(\frac{\lambda}{\mu}\right)^W \frac{P_0}{W!} = \frac{\left(\frac{\lambda}{\mu}\right)^W \frac{1}{W!}}{\left[\sum_{k=0}^{W} \left(\frac{\lambda}{\mu}\right)^k \frac{1}{k!}\right]} \qquad \ldots(5)$$

Thus, the arbitration latency $\Delta_{\text{arbitrate}}$, can be calculated using the probability transitions shown in Figure 65a and the latency of acceptance and denial on each arbiter.

$$\Delta_{\text{arbitrate}} = \frac{P_a \times L_{accp} + P_a \times P_b \times (W + L_{accp}) + P_a \times P_b^2 \times (2W + L_{accp}) +}{P_a \times P_b^3 \times (3W + L_{accp}) + P_b^4 \times (4W + 2)} \qquad \ldots(6)$$

The last section in the architecture is the scoreboard. The access through the *scoreboard* is regulated by means of a FIFO. All requests accepted by the arbiter are queued inside the scoreboard. The *scoreboard* access is given to the request stored at the FIFO head. The clock cycle latency through the *scoreboard* depends upon the waiting period in FIFO and also the scoreboard access and update time. The *scoreboard* access and update time τ (in clock cycles) is fixed however the waiting period inside the FIFO is variable and depends upon the number of requestors accepted by the arbiter previously and that have not yet accessed the *scoreboard*. The arrival at the FIFO is modeled as a Poisson process with rate Λ packets/sec dependent upon the acceptance probability P_a of the arbiter and the arrival rate λ packets/sec at arbiter $(\Lambda = \lambda P_a)$. The average wait time in FIFO is calculated using a *M/D/1* system, which has deterministic service times.

$$\Delta_{\text{scoreboard}} = \textit{(Scoreboard update latency in clock cycles) + (FIFO wait latency in clock cycles)}$$

$$\Delta_{\text{scoreboard}} = \tau + \left(\frac{\Lambda \times \frac{\tau}{CR}}{2\left(1 - \frac{\Lambda \times \tau}{CR}\right)}\right) \times CR$$

$$\Delta_{\text{scoreboard}} = \tau + \left(\frac{\lambda P_a \times \frac{\tau}{CR}}{2\left(1 - \frac{\lambda P_a \times \tau}{CR}\right)}\right) \times CR \qquad \ldots(7)$$

Figure 66   shows the average   forwarding latency   for 4-fiber router and different number of wavelengths *W*/fiber. As the load increases, the average forwarding latency, as expected, increases for different *W*/fiber. The load is function of the aggregate packet arrivals and the capacity of the router.

$$\text{Load} = \frac{\lambda}{capacity} = \frac{\lambda}{\left(\dfrac{DR \times N \times W}{average\_packet\_size}\right)}$$

The average latency also increases with the increase in number of wavelengths *W*/fiber, as *IChannels* request for *W* clock cycles at each arbiter, before attempt forwarding through other output arbiters in the contention resolution algorithm. Compared with 4 wavelengths/fiber, the router capacity experiences a 8-fold increase for 32 wavelengths/fiber while the average forwarding latency experiences a 3-fold increase. This increase in forwarding latency is easily compensated by migrating from FPGA based controller and using 0.18mm or better ASIC technology.
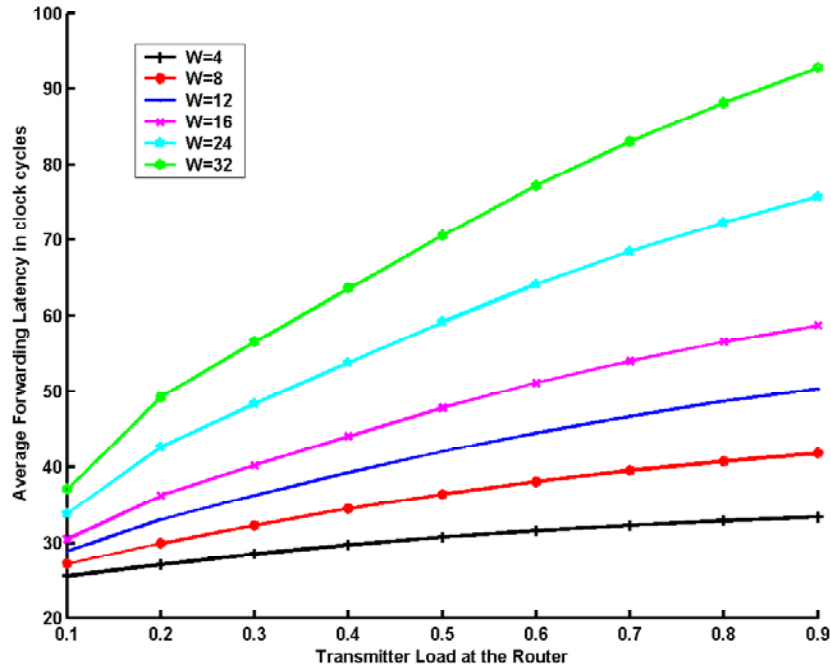
Figure 66:   Average forwarding latency in clock cycles as a function of transmitter load for 4 fibers and different wavelengths W/fiber

The forwarding latency depends upon the number of contentions that need to be resolved in order to route the packet. The hardware implemented router controller supports 4 fibers and 4 wavelengths/fiber. In hardware implementation when a packet is routed without any contentions the best-case forwarding latency of the label is 26 clock cycles, however in the worst case when the packet is dropped, as it could not be routed through wavelength, time and space domain contention resolution, the worst-case forwarding latency is 38 clock cycles. For the implemented router controller, Figure 67 shows the comparison of the observed worst-case and best-case forwarding latency with the average forwarding latency as obtained through the analytical model using equation 1. For low loads there are few packet

contentions and the average clock cycle latency is close to 26 clock cycles. However, the average latency increases with the increase in load as on an average the incoming packets experience a greater contention, but the average is around 32 clock cycles for a load of 0.9. Thus, the analytical model accurately agrees with the implemented hardware.
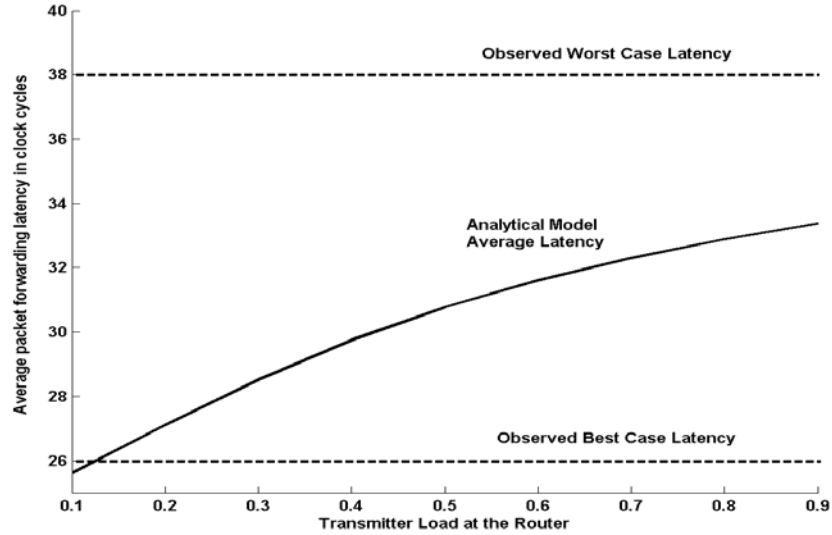
Figure 67:   Comparison between analytical and observed latency in clock cycles for the implemented 4-fibers and 4-wavelengths/fiber router controller.

5.3. Optical packet routing and switching Experiments
5.3.1 Subcarrier transmitter design

This part is mainly focused on the circuit board for 14GHz subcarrier transmitter. First, we chose the high frequency laminate from Rogers Corporation as the circuit material. According to its data sheet, this material is easy to fabricate and has very low loss to high RF frequency. Its dielectric constant is 2.94 and its dissipation factor is 0.0012. Then we chose the dielectric thickness and copper cladding as 0.010 inch and 1 oz respectively for my design. The next step is to design the RF filter on the circuit. There are two filters to be designed, one is the low pass filter with a cutoff frequency at 10GHz for the payload; the other is the band pass filter with a bandwidth of 1GHz and centered at 14GHz for the SCM label. Since we had all the parameters for the material, it is very easy to design this filters with Agilent's ADS. The following Figure 68 are the layouts and simulation results for those filters.
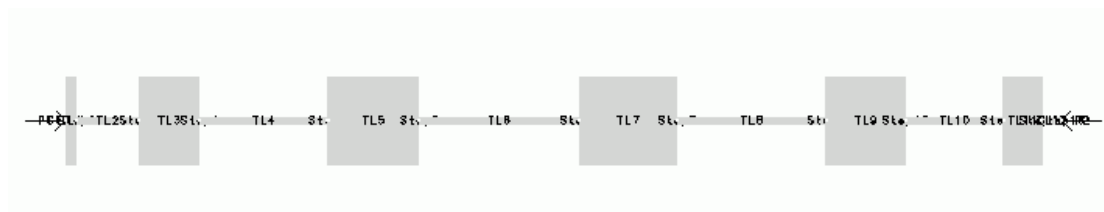
Figure 68:   Layout of LPF
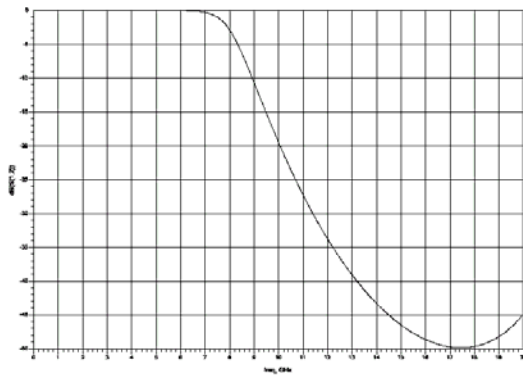
Figure 69: Frequency response of LPF
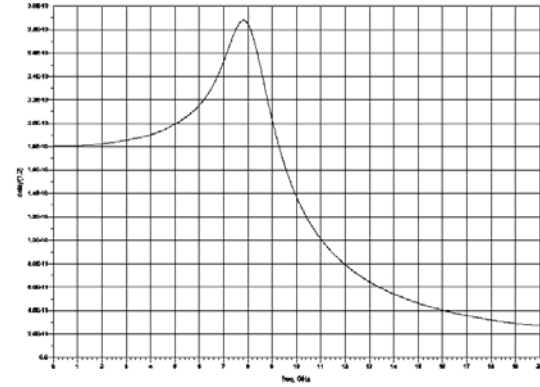
Figure 70: Group delay of LPF

Figure 71: Layout of BPF
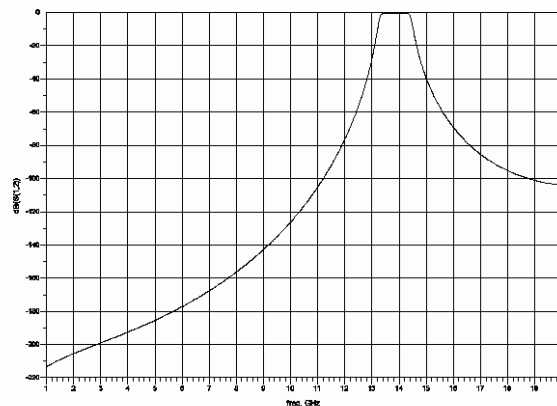
Figure 72: Frequency response of BPF

After finishing the filter design, the work is to do the layout of the circuit board. Basically, the circuit should content two parts. One is for testing individual component, such as filters, mixer and so on; the other is for functionality of the subcarrier modulation and transmission. The layout of the whole circuit board is below:

64
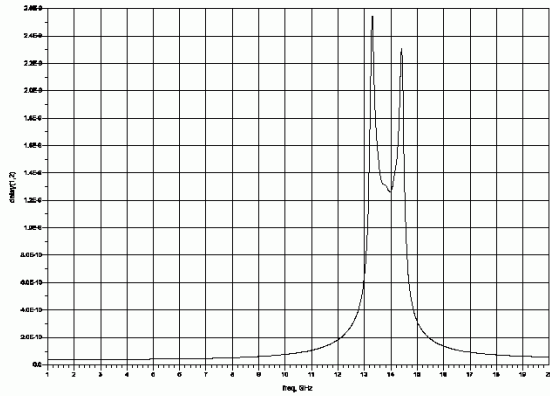
Figure 73: Group delay of BPF Fig.
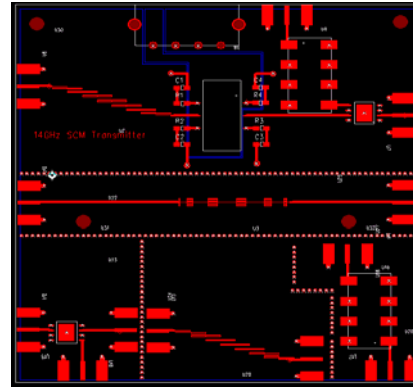
Figure 74: Circuit Layout of SCM transmitter

As shown in the Figure 74, the upper two sections are the functional sections and the lower three sections are for testing. The top section is for subcarrier modulation of the label and the second one is for payload, only two SMA connector and one low pass filter.

After finishing the circuit layout, we sent out the board for fabrication. And we have already received the boards. Now we are doing the testing of the board. The band pass filter is OK with the bandwidth, but the center frequency of it shifts a little. Now the center frequency is 13.3GHz, but not 14GHz. In our opinion, this shift is caused by the fabrication error. Since the traces for BPF are very small, a very small error will change the filter's characteristics a lot. For low pass filter, the response is OK. In the future, we will do the testing for the mixer and the amplifier. Another shortcoming of the circuit board we should point out here is that it is too fragile to be handled. When it is broken, its loss will increase and the frequency response will change. The really circuit for the transmitter is below:
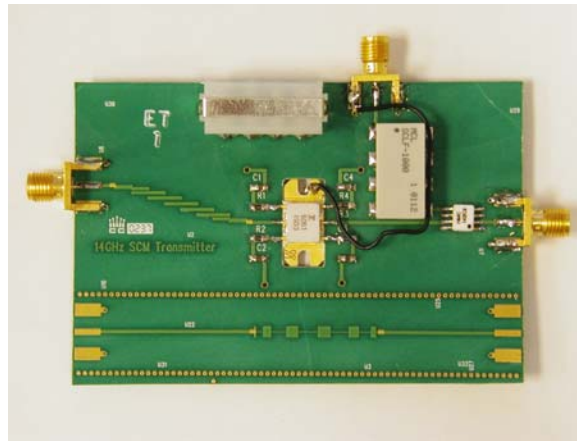
Figure 75: Real circuit of SCM transmitter

Another thing we did in this summer is testing and setting up the EA modulator for subcarrier transmission. We setup the SCM transmitter with EA modulator. But when we compare the result with the signal modulated by LiNbO$_3$ modulator, it seems that the frequency response of the EA modulator is not so good as LiNbO$_3$ modulator in high frequency from 8 to 15 GHz. The following is the optical spectrum of the signals when the subcarrier frequency is 12GHz:

<div align="center">(a)   (b)</div>
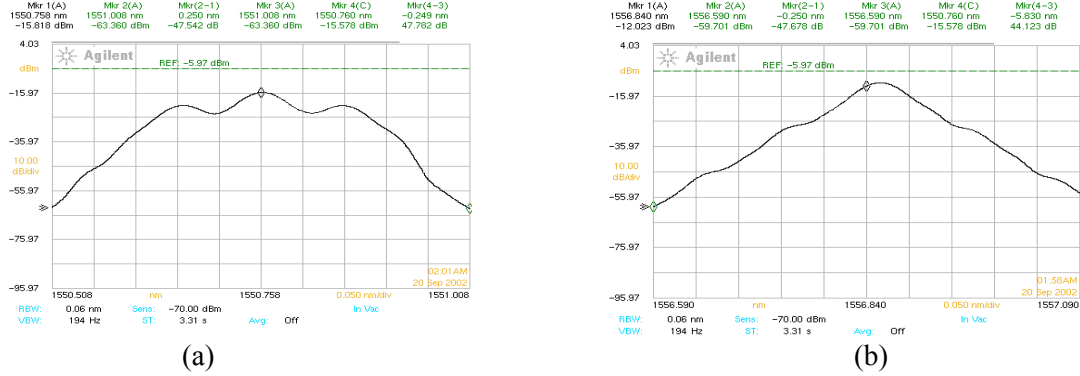
Figure 76: Optical spectrum of SCM label and payload modulated by
(a) LiNbO$_3$ modulator (b) EA modulator

In the future, we will continue doing testing for my circuit board and also for the EA modulator. We will try to extract the label signal with FBG and optimize the output signal.

5.3.2 Tunable laser driver and control circuit board

5.3.2.1 Introduction

The progress in the tunable laser driver and control circuit board includes a better understanding of the calibration data, which was obtained from the first version board, and the assembly and debug of the second version. Simulation work on tunable laser is also proposed. The following sections will talk about the above topics one by one.

As a brief review, in the optical-label switching router architecture, the function of the tunable laser and driver is to provide laser source at a desired wavelength for the wavelength converter. Combined with the wavelength converter they form a tunable wavelength converter subsystem. As shown in Figure 77, the tunable laser and driver receives electrical control signal from the FPGA, set the wavelength accordingly and output continuous wave light to the wavelength converter. The basic design of the driver is as following: The tunable laser consists of 4 sections – gain, front grating, rear grating, and phase. The tunable laser driver circuit board sends a constant current to the gain section to maintain laser power. It converts the digital control signal from the FPGA into analog current outputs to the other 3 sections using high-speed DACs, thus controls the lasing wavelength. The requirements for the tunable laser driver and control board are mainly imposed by the components after it, especially the wavelength converter, the arrayed waveguide grating (AWG; a static wavelength switch) and the optical network outside the optical router. The key requirements are stability in power and wavelength, accuracy in wavelength, and high speed in wavelength tuning.

Figure 77:   Tunable laser and driver in the architecture

5.3.2.2 The first version of the tunable laser driver and control board

Figure 78   shows the   block diagram of the tunable laser driver and control circuit.   The D/A converters AD9731 control each current going into the front, rear or phase section. The input of AD9731 is the digital data of the current value. Because FPGA has limited output ports, it provides the most significant 4 bits of the data for front and rear currents. Dipswitches on the board control all the remaining bits and the phase current. Controlling the most significant 4 bits of both the front and the rear grating currents data is adequate to access the useful wavelengths for most of the experiments when channel number is limited to 3. Sometimes the wavelength is not accurate or unwanted side-lobes show up. Then the dipswitches will be adjusted to optimize the laser output.
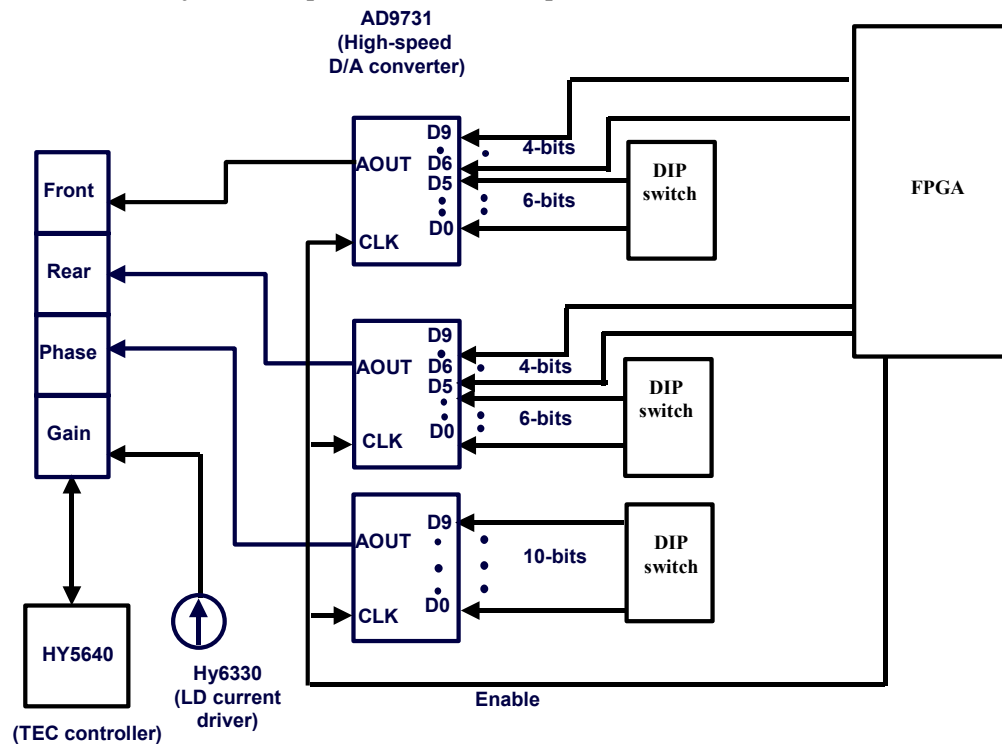
Figure 78:   Block diagram of the tunable laser driver and control circuit

67

According to the block diagram the circuit is designed. The PCB layout and manufacturing were contracted to a PCB fabrication company. The assembly and testing of the board was done in our lab. Figure 79 is a picture of the finished board.
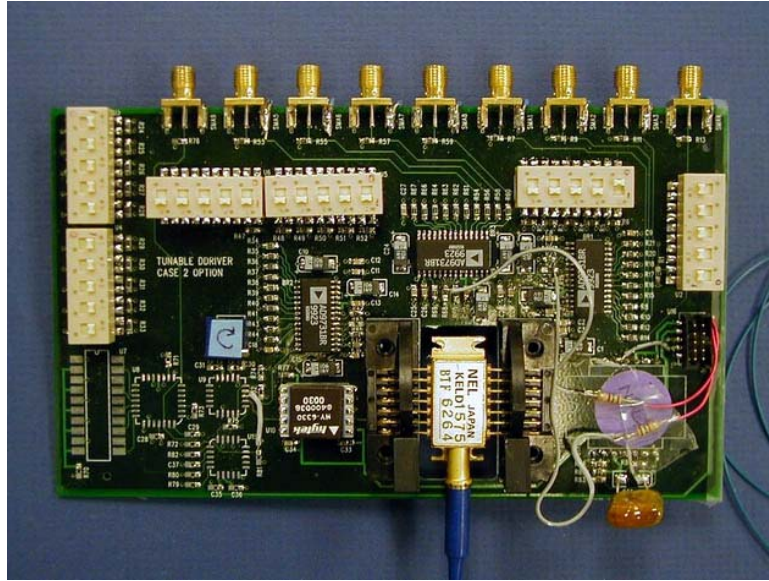
Figure 79:   The finished tunable laser driver and control board version 1, with the tunable laser installed.

The wavelength switching characteristics of the tunable laser were tested. For static characteristics, 256 data points were measured completely to cover all combinations of the most significant 4 bits of both front and rear DAC. Figure 80 and Figure 81 show the static characteristics (wavelength versus grating currents).   Figure 82 shows the wavelength coverage (without adjusting the phase current or the remaining less significant bits in front and rear grating currents). Most of the wavelengths in the range of 1540nm to 1570nm can be accessed. A few hops exist, the wavelengths in which can be covered in the future by adjusting the phase current.

Figure 80:   Static tuning characteristics.
(a) Wavelength v.s. front grating current (rear grating current is also changing).
(b) Wavelength v.s. rear grating current (front grating current is also changing).

Figure 81:   Wavelength vs. both front and rear digital bits; note that the larger the digital data input, the smaller the real current
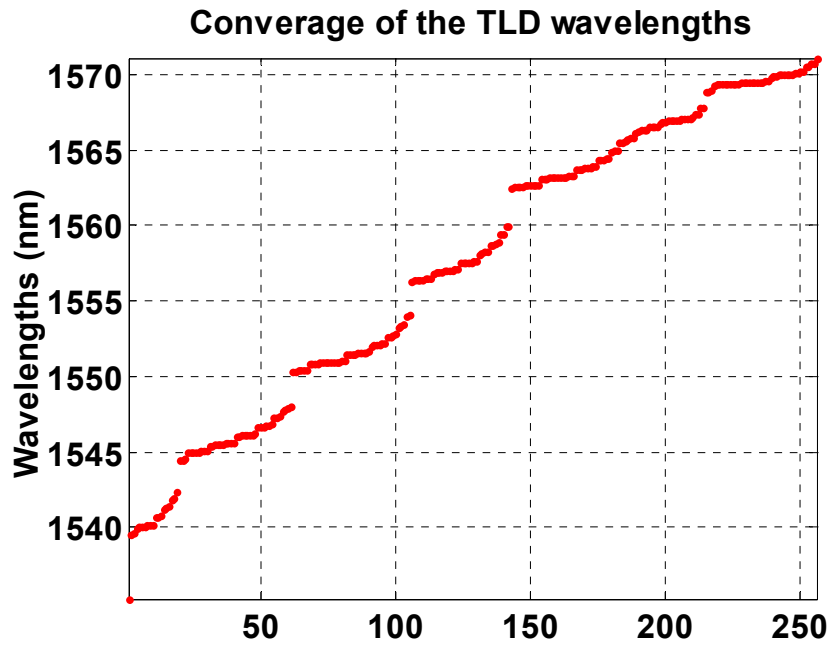
Figure 82:   Wavelength coverage of the tunable laser

To obtain the optical spectrum and rise time of the laser output under dynamic switching, experiment setups shown in Figure 83 were used. The results are shown in Figure 84. All results showed clean peaks and short rise time (~1ns).

	(a)		(b)
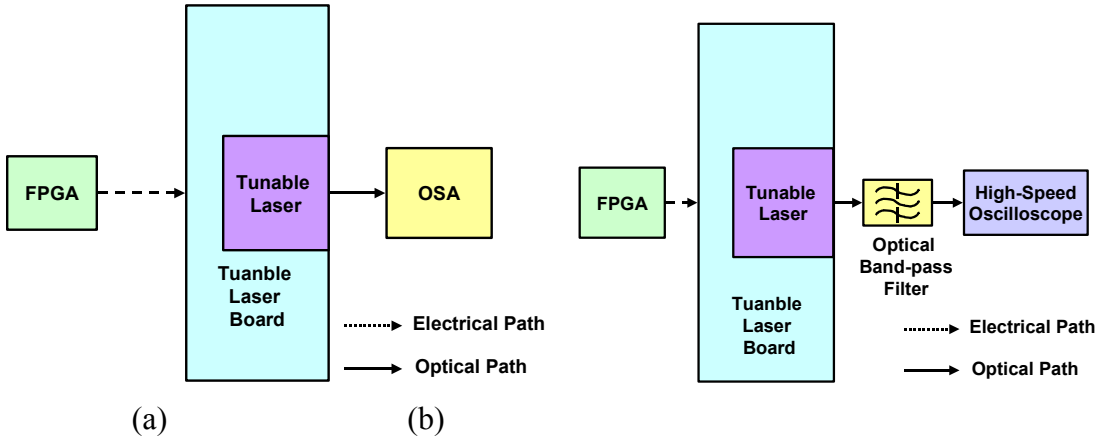
Figure 83: Experimental setups. (a) To obtain the spectrum under switching. OSA stands for optical spectrum analyzer. (b) To obtain switching rise time (the oscilloscope is Agilent Infinium DCA 8610A which has optical input).

	(a)		(b)
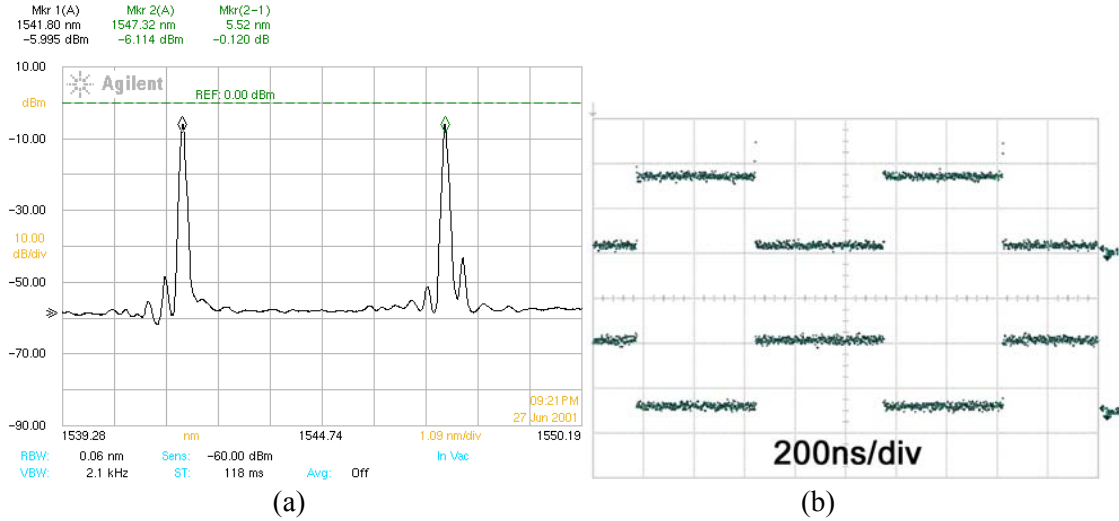
Figure 84: Experimental results.
Spectrum of two-wavelength switching. The two wavelengths are 1541.80nm and 1547.32nm.
Oscilloscope showing the fast switching of wavelength (after optical filter). The top waveform is for one wavelength while the bottom for the other.

Discussed above is the simplest design to realize the necessary function of the driver and control circuit. The design does not have any logic functionalities. Thus it depends on the FPGA to memorize the mapping of digital data to actual currents. This scheme works fine for a few channels, but it does not scale. The better way is to store the mapping information locally in the tunable laser driver and control board. Because of the difference among individual lasers, this mapping will not be exactly the same for each laser even if they are the same model from the same manufacturer. Thus a calibration program is also necessary on the tunable laser driver and control board to find out the specific mapping for the laser installed and save them for future use. Another functionality that is necessary is the wavelength and power stabilization circuit.

5.3.2.3 The second version of the tunable laser driver and control board

The most important functionality that has been designed into the second version of the board is the table of the mapping from digital data to real current. For this purpose, an Static Random Access Memory (SRAM) and a micro-controller are introduced. This relieves the FPGA of tracking the current-wavelength relationships for individual tunable lasers. Instead each relationship is stored in the form of a look-up table and burned into the micro-controller. On start-up, the table will be loaded into the SRAM. Thus when FPGA sends the "code" for a particular wavelength, the corresponding digital current values will be searched and found in the SRAM. The DACs will then convert the digital value into analog current to tune the wavelength of the tunable laser. Also, an RS232 serial port is provided for the connection to a computer. The computer can communicate with the micro-controller to perform more sophisticated tasks, for example automated tunable laser calibration. A method to avoid unwanted intermediate wavelength from showing up is also attempted. The method requires turning off the tunable laser before grating currents are changing, and turning it on again after the grating currents are stable. Two variable delays are introduced to control the timing. However, the low-speed characteristics of the current driver does not allow this design to work. Therefore, that part of the circuit is not used in the realized board. Figure 85 shows the block diagram and Figure 86 shows the finished board. Measurement of the static wavelength tuning similar to the one discussed in first version is applied. Figure 87 to 89 show the results.

To demonstrate the capability of the new board, we repeated the contention resolution experiment in wavelength, space and time domain using optical label-switching. Two second version tunable laser driver and control boards are used. First, wavelengths around desired AWG channel wavelengths are located using the calibration data. Second, less significant bits and phase current are adjusted to match the AWG channel wavelength precisely. The digital values of the front, rear grating currents and the phase current are recorded. Third, the micro-controller is programmed so that these digital values will be loaded into the SRAM upon power-on.

Figure 85:   The block diagram of the second version tunable laser driver and control circuit.

71

Figure 86:   The second version of the tunable laser driver and control board.
(a) The realized board. (b) The board working (the OSA shows the tunable laser output).

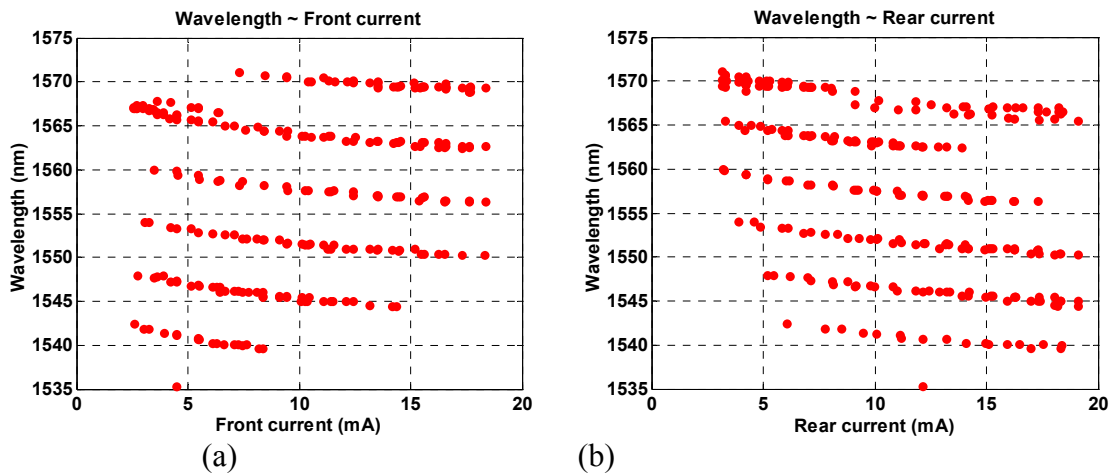Figure 87:   Static tuning characteristics.
(a) Wavelength v.s. front grating current (rear grating current is also changing).
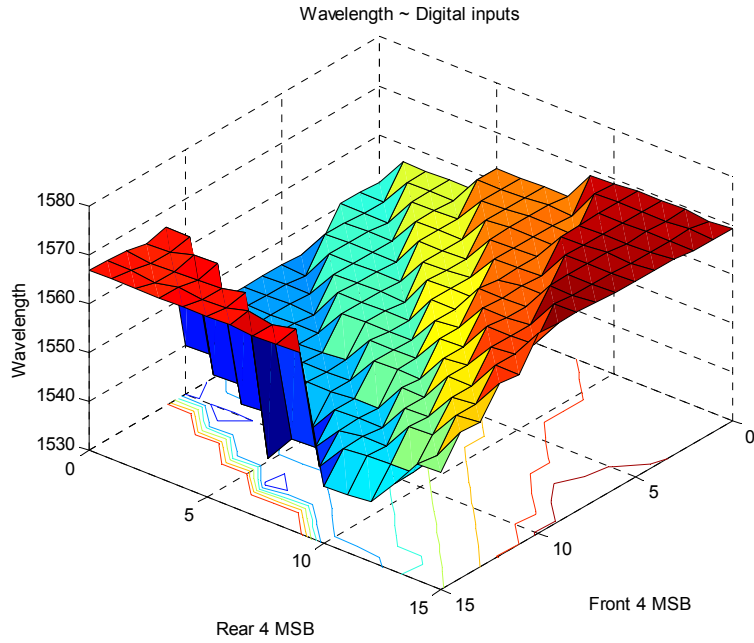(b) Wavelength v.s. rear grating current (front grating current is also changing).

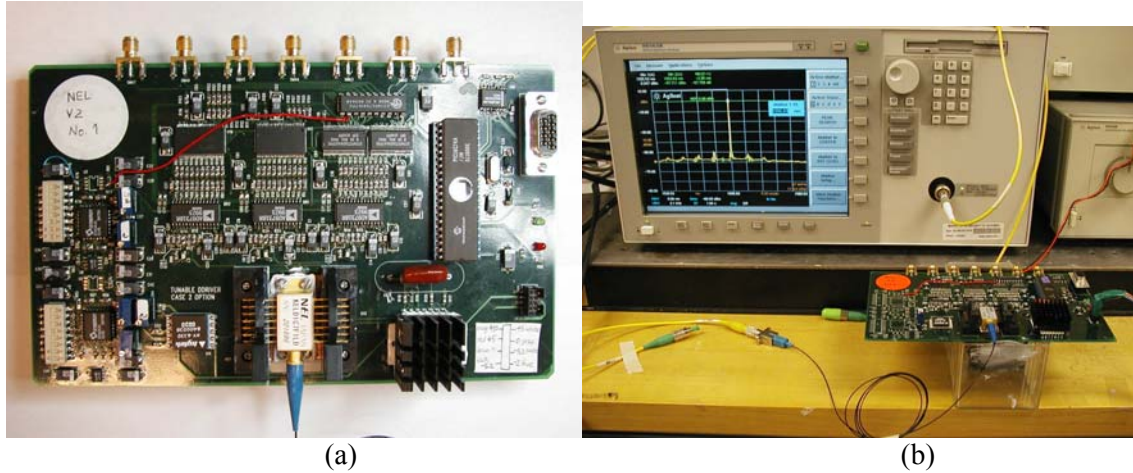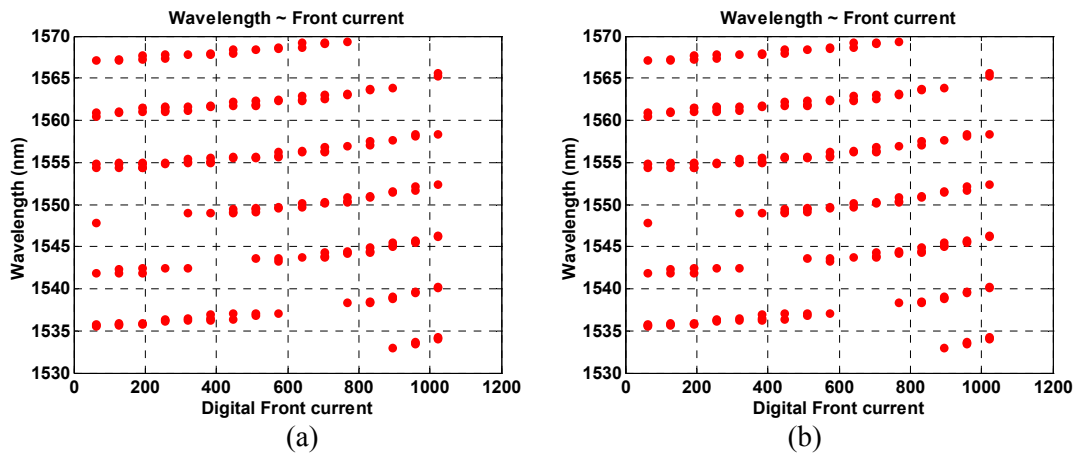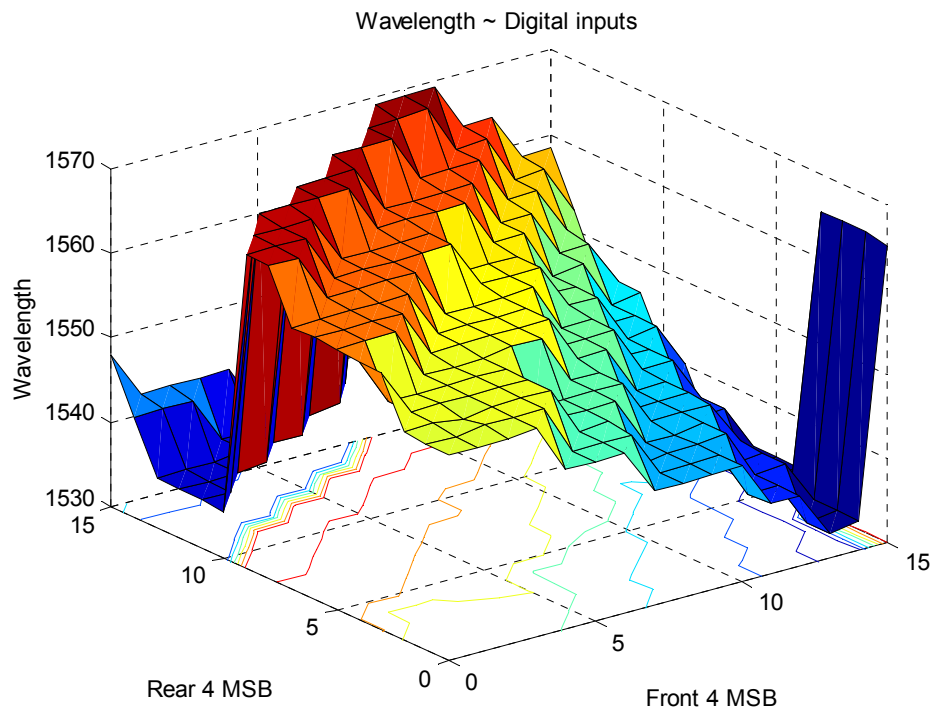Figure 88: Wavelength vs. both front and rear digital bits; note that the larger the digital data input, the smaller the real current
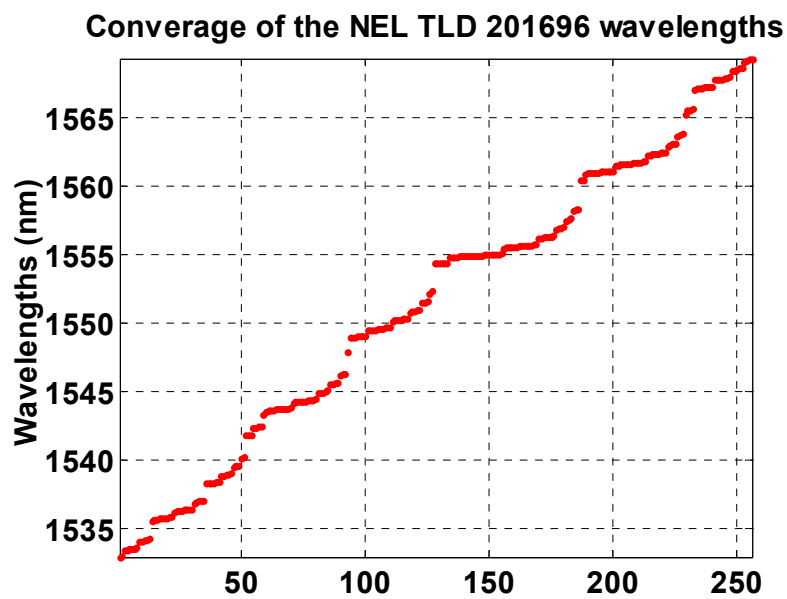

Figure 89: Wavelength coverage of the tunable laser

5.3.3 All-optical packet switching based on label-swapping system with 2R regeneration

5.3.3.1 Introduction

Label-swapping technology is important for providing scalability in multi-protocol label switching (MPLS) and optical-label switching (OLS) networks. Optical sub-carrier multiplexing (SCM) is an attractive label-encoding scheme for accommodating both the sub-carrier label and the baseband data payload on the same optical wavelength channel but on separate modulation frequencies. Compared to bit-serial time-domain label-encoding techniques, SCM techniques facilitate the label swapping based on frequency-dependent separation of the label and the data payload. Optical-label swapping techniques reported to date involve relatively complex single side-band transmitters [1], nonlinear optical schemes with bulky fiber spools [2], or overmodulating schemes with inter-modulation penalty [3]. This report proposes and demonstrates an all-optical packet switching based on label-swapping technique with 2R regeneration and no inter-modulation effect. The experiment shows error-free operation with negative power penalty.

5.3.3.2 Optical label swapping Experiments

Figure 90 shows the experimental setup for regenerative all-optical l abel swapping. It consists of two stages: the first stage performs label extraction of the incoming optical packet. The second stage performs label rewriting and 2R regeneration for data payload. Inside the left dashed-line box of Figure 90 is a transmission line module consisting of two erbium-doped fiber amplifiers (EDFAs), one optical attenuator and a 75km SMF-28 fiber acting as the noise and dispersive element to test the regeneration capability of the label-swapping module. The parallel bit error rate tester (ParBERT) synchronously generates the electrical label and data payload signals. The SCM transmitter generates the optical packets using double-side band (DSB) SCM with a 14 GHz sub-carrier. A packet consists of a set of time-overlapping label and data payload. The label bit rate is 155 Mb/s while the data payload bit rate is 2.488Gb/s. The packet duration is 619.4 ns (1536 bit payload and 96 bit label), with 206.5 ns guard time between packets. The data payload contains a 2^{15}-1 pseudo-random bit sequence (PRBS). The fiber Bragg grating 1 (FBG 1) has a sharp (~10 GHz FWHM) high reflectivity band (>99.9%) peaking at the same wavelength (1550.75 nm) as the wavelength of the SCM transmitter. Therefore, the FBG1 reflects the baseband data payload signal and transmits the label signal, thus achieving all-optical separation and extraction of the label in label extractor 1 [4]. The transmitted label signal reaches the optoelectronic receiver for bit-error-rate (BER) performance measurements of the label. At the same time, the reflected data payload signal goes to the label-rewriting module after being amplified by an EDFA.
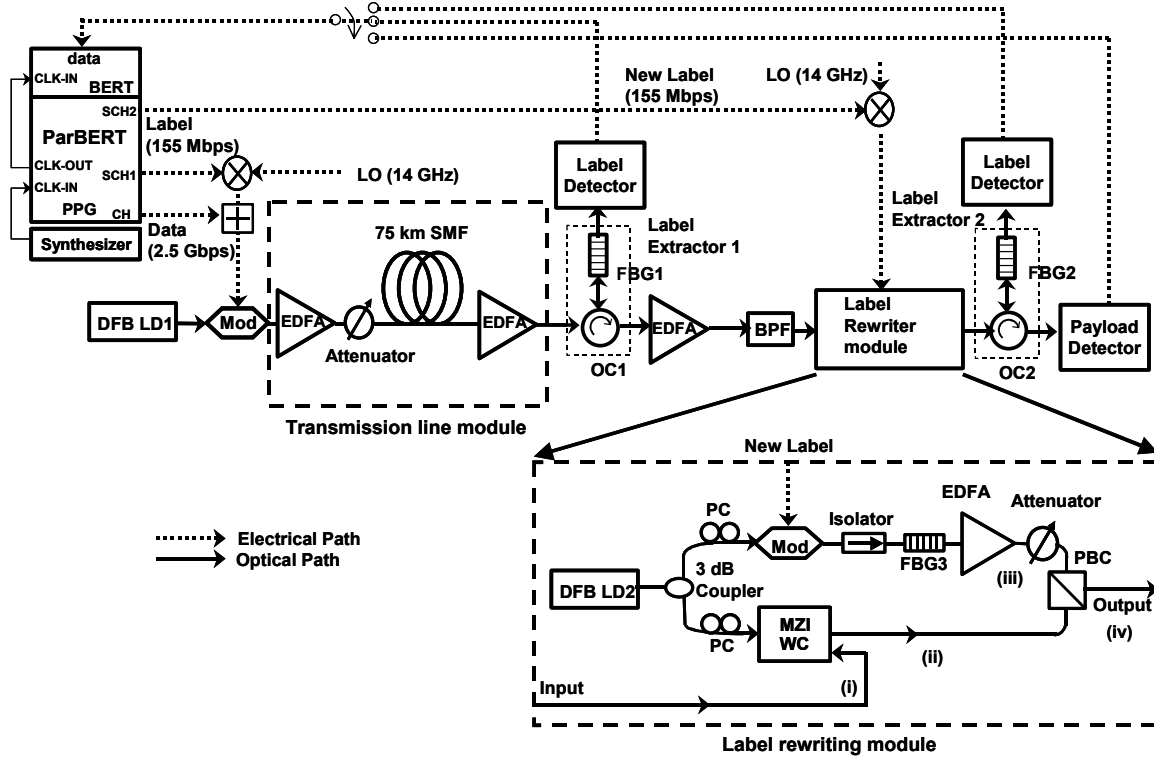
Figure 90: Schematic diagram of the experimental setup for regenerative all-optical label swapping (Par BERT: Parallel Bit Error Rate Tester; LO: Local Oscillator; EDFA: Erbium Doped Fiber Amplifier; Mod: LiNbO$_3$ Modulator; FBG: Fiber Bragg Grating; BPF: Band Pass Filter; OC: Optical Circulator PC: Polarization Controller; MZI WC: Mach Zehnder Interferometer Wavelength Converter; PBC: Polarization Beam Combiner)

The dashed-line box at the bottom right of Figure 90 illustrates the schematic diagram of the optical-label rewriting module. It consists of a distributed feed-back laser diode (DFB LD2), a LiNbO$_3$ modulator, a 1x2 fiber coupler, two polarization controllers (PCs), a semiconductor optical amplifier based Mach-Zehnder interferometer wavelength converter (MZI WC, Alcatel 1901-ICM), an isolator, a FBG, an EDFA, an attenuator, and a polarization beam combiner (PBC). This setup achieves 2R regeneration of the data payload on one arm and sub-carrier modulation of the new label on the other arm, both onto a new wavelength. The DFB LD2 provides continuous-wave (cw) light to both arms through a 1x2 fiber coupler. The center wavelength of the DFB LD2 inside the label-rewriting module is 1555.73 nm. 50% of the cw emission travels to the SOA-based MZI WC to be used as a probe beam for wavelength conversion by cross-phase modulation driven by the data payload signal obtained by removing the label from the incoming packet in label extractor 1. The other 50% of the cw emission is modulated by the SCM signal containing the new label information generated by the switch controller (in this case Par BERT). In order to avoid interference with the data payload and the unwanted baseband cw component of the SCM, the setup includes FBG3 with peak reflectivity centered at 1555.73 nm, rejecting the baseband signal and allowing only the DSB SCM label signal to transmit. The isolator prevents feedback into the DFB LD2. The PBC ensures that the two optical signals from the two arms are combined together without any undesired coherent interference. The two PCs are adjusted to achieve high throughput and optimum performance.

Figure 91 (i) and (ii) show the optical spectra measured at the input port (i) and at the output port (ii) of the MZI WC respectively. They contain only the baseband data payload. Figure 91 (iii) shows the optical spectrum of the new label without baseband cw component measured at port (iii). Figure 91 (iv) shows the final output spectrum measured at port (iv), and the result is a combination of the baseband

signal containing the data payload and the SCM signal containing the new label. This combined signal is sent to label extractor 2, which has the identical structure as label extractor 1 except for the FBG2 peak wavelength (1555.73 nm) in order to achieve the BER measurement on the new label and the regenerated data.
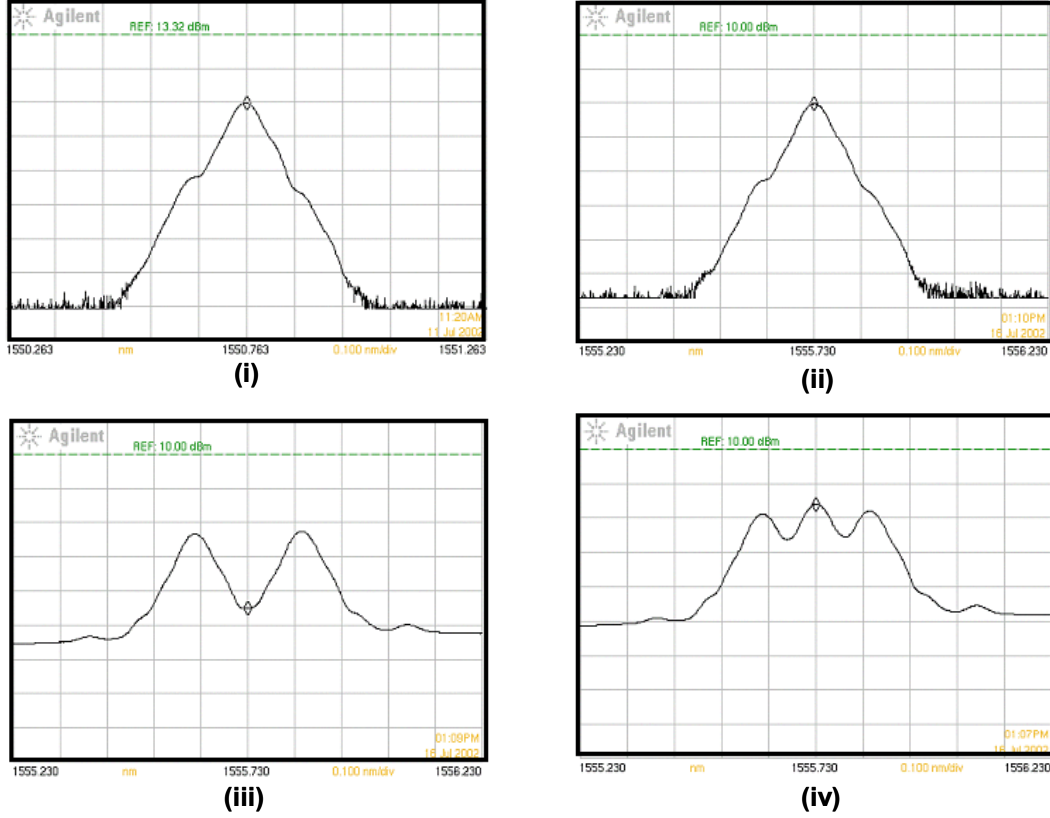

Figure 91: Optical spectra of (i) input payload (1550.76 nm), (ii) regenerated payload (1555.73 nm), (iii) new label (DSB_SCM) without baseband, (iv) regenerated DSB-SCM (0.1 nm/div)

Two sets of accumulated packet-by-packet BER measurements were carried out for the setup with and without the transmission-line module shown in Figure 90. Figure 92 (a) shows the BER and eye diagrams of the label and payload without the fiber-transmission-line module. The label-swapping system imposes essentially no power penalty on either the label or the payload. Figure 92(b) shows the results for the second set of measurements where the transmission-line module was included to purposely inject deteriorated label and data payload signals to test the regeneration capability. The 25 dB attenuation in addition to the 75 km SMF fiber transmission followed by the EDFA resulted in the degraded extinction ratio of 8 dB for the data payload, which improved to greater than 13 dB after the 2R regeneration. As Figure 92 (b) indicates, the label-swapping system achieved a 3dB negative power penalty for the data payload and a 2.5 dB negative power penalty for the label for this set of measurement. The 2R regeneration effect for the data and the label rewriting effect for the label are clear in this measurement.
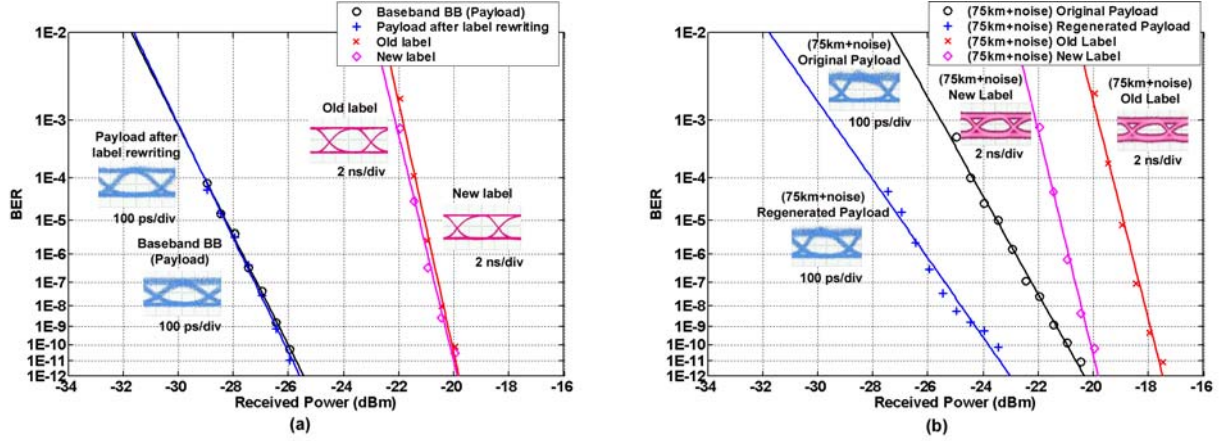
Figure 92: BER test results for label swapping (a) without transmission-line module, (b) with transmission-line module (Insets: eye diagrams of the label and payload before and after label swapping)

5.3.3.3 Optical packet switching based on optical label swapping

The experimental set up for all optical packet switching based on the optical label swapping is shown in Figure 93. It consists of an optical SCM Tx, two label exctractor consisting of a circulator and a FBG, a header detector as a burst mode receiver, a FPGA, a tunable wavelength converter consisting of a tunable laser and a SOA, a label rewriting module and an uniform-loss-cyclic frequency (ULCF) arrayed-waveguide grating [5]. A parallel bit error rate tester (ParBERT) creates a 2.5 Gbps signal alternating between two 1536-bit packets, called Payload 1 and Payload 2. Simultaneously, the ParBERT generates a 155 Mbps signal consisting of two alternating 32-bit packets called Label 1 and Label 2. A packet consists of time-overlapping label and payload. The Burst mode receiver detects the 96 bits long optical-label herder signal, the forwarding table compares against the content of the table, and the controller generates a switch control signal and a new-optical-label content. The switch control signal from the controller is sent to the rapidly tunable laser, instructing it to tune to the wavelength corresponding to the desired output destination port of the AWGR. The experiment utilized two optical-label contents for two outputs wavelength 1 (1546 nm) and 2 (1552 nm). The data payload is delayed via the 80 m EDFA including an attenuator providing approximately 400 nsec optical delay. This data payload is input to SOA whose gain will be modulated by the amplitude modulation on the data payload. This gain modulation will be probed by the counter-propagating output of the tunable laser, which will now assume the inverted modulation information of the data payload. The bit-error rate (BER) was measured for the back-to back and after switching through the label rewriting. Figure 94 shows the bit-error rates for the optical label switching router vs. back-to-back, respectively. Both outputs show error free operations. The power penalty @ BER = 1E-9 were about 3.5 dB. This is due to the logic inversion after cross gain modulation from the SOA wavelength converter. Insets in Figure 94 show the eye diagram of the optical back to back and after OLSR. Figure 95 shows the optical spectra after switching. The optical spectrum after wavelength conversion is shown in Figure 95 (a). The switched wavelengths are 1546 nm and 1552 nm. Figure 95 (b) shows the optical spectrum after OLSR.
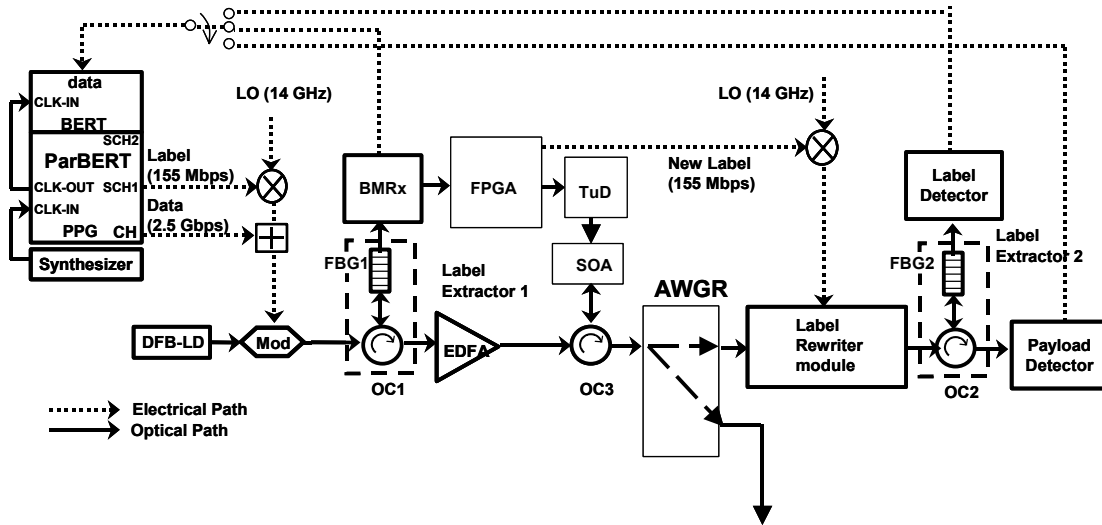
77

Figure 93: Schematic diagram of the experimental setup for optical packet switching
(Par BERT: Parallel Bit Error Rate Tester; LO: Local Oscillator; EDFA: Erbium Doped Fiber Amplifier;
 Mod: LiNbO $_3$ Modulator; FBG: Fiber Bragg Grating; OC: Optical Circulator; AWGR:Arrayed
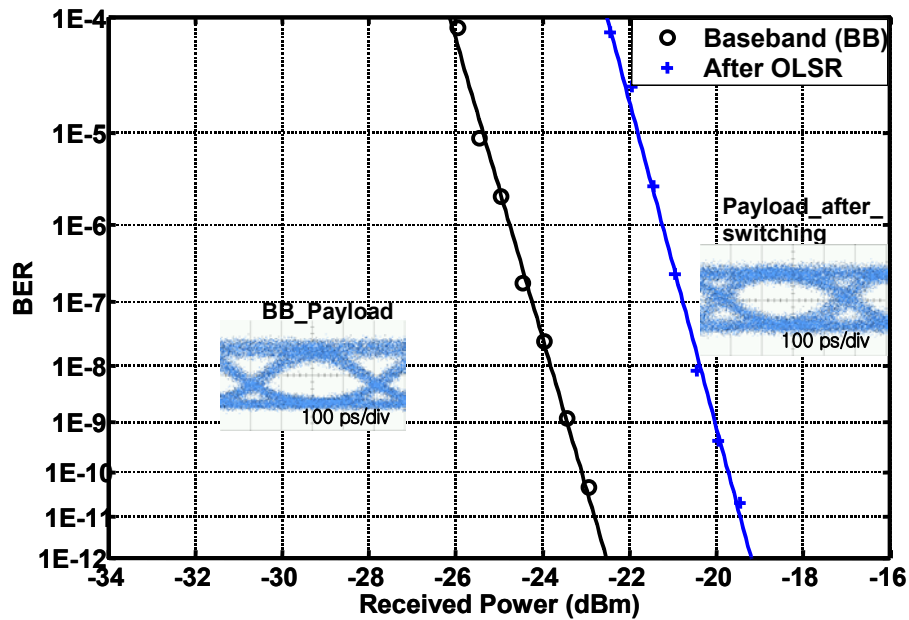Waveguide Grating Router; TuD : Tunable Laser Diode; SOA : Semiconductor Optical Amplifier)

Figure 94: BER test results for optical packet switching
(Insets: eye diagrams of the payload before and after packet switching)
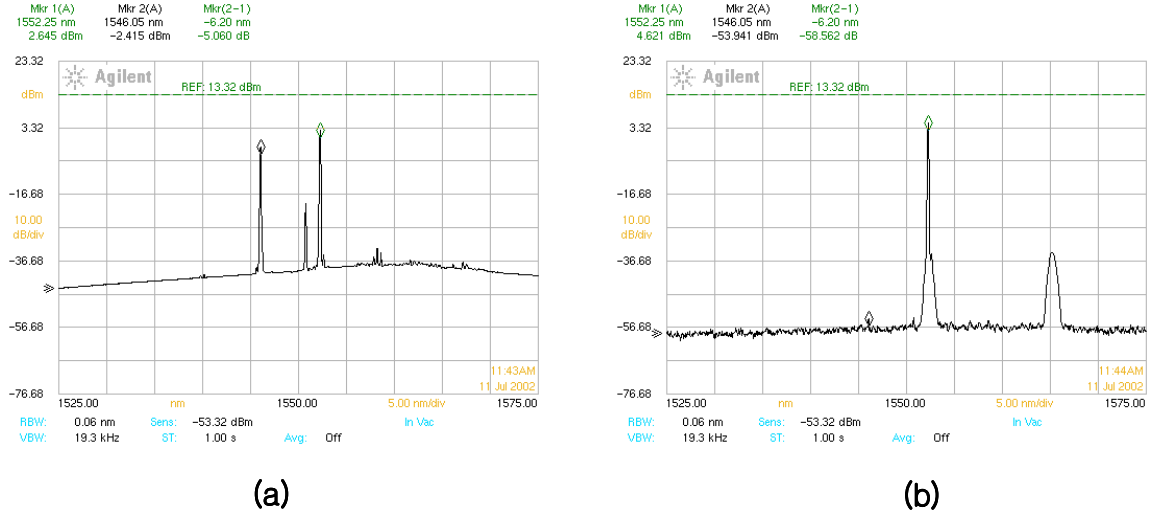
(a)   (b)

Figure 95:   Optical spectra of (a) after wavelength conversion (1546 nm and 1552 nm),
(b) after OLSR (1552 nm)

5.3.3.4 Summary

We proposed and demonstrated an all-optical packet switching based on label-swapping system with 2R regeneration and wavelength conversion. The all-optical label-swapping system achieves zero penalty for a high-quality input signal, and negative penalty for a deteriorated input signal for both labels and data payloads. The error-free, regenerating capabilities of the demonstrated all-optical label-swapping system imply its application in a scalable optical-label switching network

References

[1] Y. M. Lin, W. I. Way and G. K. Chang, "A novel optical label swapping technique using erasable optical single-sideband subcarrier label", IEEE Photon. Tech. Lett. **12**, pp.1088-1090 (2000).
[2] B. E. Ollsson, P. Ohlen, L. Rau, G. Rossi, O. Jerphagnon, R. Doshi, D. S. Humphries, D. J. Blumenthal, V. Kaman and J. E. Bowers, "Wavelength routing of 40 Gbit/s packets with 2.5 Gbit/s header erasure/rewriting using all-fiber wavelength converter", IEE Electron. Lett. **36**, pp. 345-347 (2000).
[3] H. J. Lee, S. J. B. Yoo, V. K. Tsui and S. K. Fong, "A simple all-optical label detection and swapping technique incorporating a fiber Bragg grating filter", IEEE Photon. Tech. Lett. **13**, pp. 635-637 (2001).
[4] H. J. Lee, V. Hernandez, V. K. Tsui and S. J. B. Yoo, "Simple, polarization-independent, and dispersion-insensitive SCM signal extraction technique for optical switching systems applications", IEE Electron. Lett. **37**, pp. 1240-1241 (2001)
[5] S. J. B. Yoo, Hyuek Jae Lee, Yanda Zhang, S. K. H. Fong, Vincent Hernandez, Vincent K. Tsui, Srikanth Vaidianathan, Katsunari Okamoto, Shin Kamei, "High-Speed All-Optical Packet Routing System with Optical-Label Switching and Optical-Label Swapping using Rapidly Tunable Wavelength Conversion and a Uniform-Loss Cyclic Frequency AWGR," IEEE photonics technology letters, **14**, no. 8, pp. 1211-1213 (2002.)

### 5.3.4 Multi-hop cascaded operation of an all-optical label routing system
5.3.4.1 Experiment setup and results

The experiment emulates a network with multiple optical-label switching routers, each providing label-based packet forwarding.   Figure 96(a) shows an emulated optical-label switching network consisting of several optical-label switching routers (OLSR). Three types of packets, P1, P2, and P3 with labels L1,

L2, and L3 respectively, ingress into the optical label switching networks. The first OLS router (OLSR1) performs the optical-label based forwarding of all three packets. OLSR1 forwards P3 north to a neighboring OLS router (OLSR2) and forwards P1 and P2 east to another neighboring OLS router (OLSR3). The OLSR3 in turn forwards P1 and P2 to two different output ports. For multi-hop scalable OLSR operations, data payload regeneration and label swapping/regeneration are beneficial.
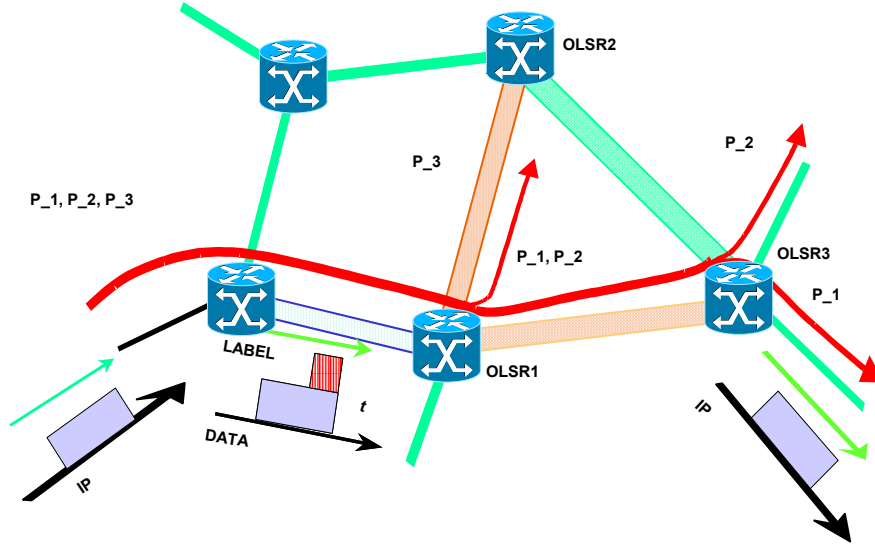
(a) A schematic of the emulated network
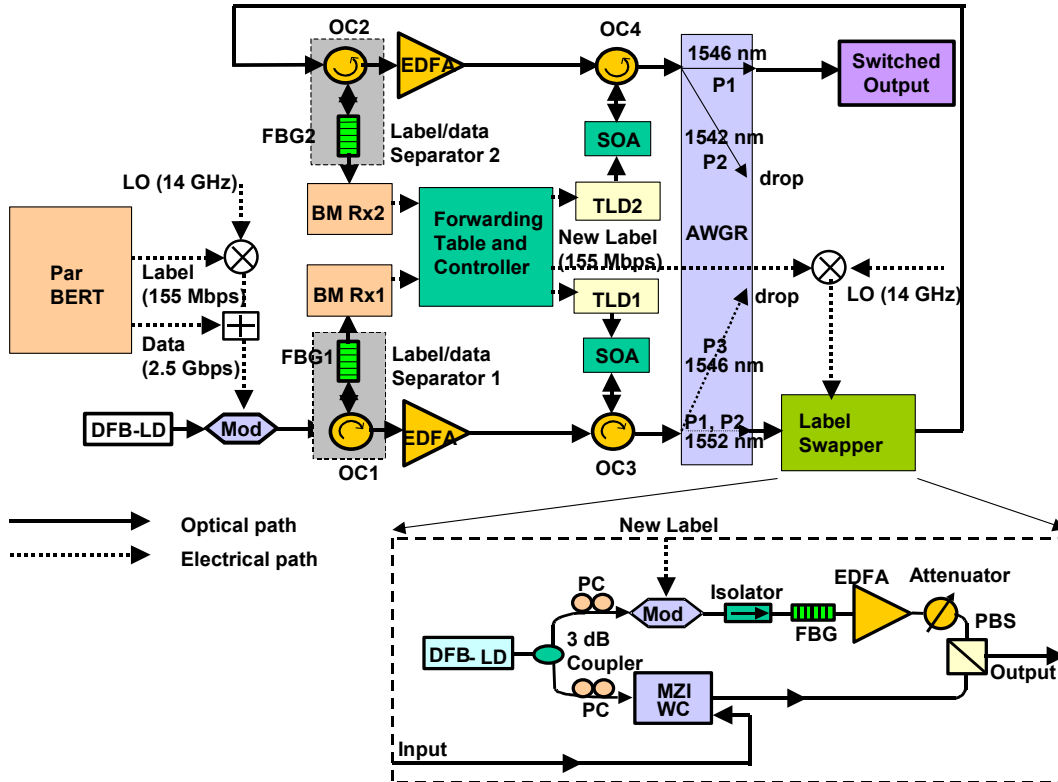
(b) An experimental setup for the multi-hop emulation of this network

Figure 96: Experimental Setup for cascaded operation of OLS system

Acronyms in figure 96 represent the following:

ParBERT: Parallel Bit Error Rate Tester; PPG: Pattern Generator, LO: Local Oscillator; Mod: Modulator; FBG: Fiber Bragg Grating; BPF: Band Pass Filter; OC: Optical Circulator; BM Rx: Burst Mode Receiver; SOA: Semiconductor Optical Amplifier; MZI WC: Mach Zehnder Interferometer Wavelength Converter; TLD:Tunable Laser Diode; AWGR: Arrayed Waveguide Grating Router; PC: Polarization Controller; PBS: Polarization Beam Splitter.

Figure 96(b) shows the detailed structure of OLSR and setup for this network emulation. As Figure 96(b) shows, the actual experiment places OLSR1 and OLSR3 on the same optical router system with multiple line cards and replaces OLSR2 with a drop port. The OLSR system consists of an optical-subcarrier multiplexing transmitter (SCM Tx), two optical label/data separators, two burst mode receivers (BMRx1 and BMRx2) for label detection, a field programmable gate array (FPGA) that implements the forwarding table and switching control, two tunable wavelength converters consisting of tunable lasers and semiconductor optical amplifiers (SOAs), a uniform-loss-cyclic frequency (ULCF) arrayed waveguide grating router (AWGR), a label rewriting modul and data receivers. The Parallel Bit Error Rate Tester/Pattern generator (ParBERT) synchronously generates the electrical label and payload signals. The LiNbO$_3$ external modulator modulates the continuous wave (CW) light from the distributed feedback laser diode (DFB LD) using a subcarrier multiplexed signal consisting of a baseband 2.5 Gb/s data payload and a 155 Mb/s label modulated onto a 14 GHz subcarrier. Hence, the modulated signal includes a double-sideband subcarrier label 14 GHz away from the center optical carrier frequency. The combination of a fiber Bragg grating and an optical circulator achieves all-optical label extraction. The BMRx asynchronously recovers the label contents from optical domain to electrical domain. The recovered label signal induces the forwarding decision inside the switch controller according to the routing algorithm in the FPGA. Based on the forwarding decision, the switch controller sends a control signal to the tunable laser (TLD) to switch to the designated wavelength. The TLD generates a tunable probe light for the SOA1, which modulates the payload signals onto the new wavelength by cross-gain modulation. Payloads with different labels are converted onto different wavelengths corresponding to the desired output ports of the AWGR.

The ParBERT generates repeated patterns of packet 1 (P1), packet 2 (P2), and packet 3 (P3) with different labels (L1, L2, L3). Figure 97(a) shows the three packets. The optical labels L1, L2 and L3 cause the wavelength of the TLD1 to be switched to λ1 (1552nm), λ1 (1552nm), and λ2 (1546 nm) respectively according to the forwarding table. The tunable wavelength converter will convert P1 and P2 to wavelength 1552 nm and P3 to 1546 nm according to the optical-label based forwarding decision. After routing through the AWGR, P3 will be dropped and P1 and P2 go to the label-swapping module as shown in Figure 97 (b). This represents the OLSR1 in the emulated network. The switch controller generates new labels for payload P1 and P2 and drives the modulator inside the label-rewriting module with a 14 GHz carrier frequency. At the same time, payload P1 and P2 will be regenerated to the fixed wavelength (1555.7nm) in the SOA Mach-Zehnder Interferometer wavelength converter (SOA-MZI WC). The packets with the regenerated label and payload transmit to label / data separator 2. The BMRx2 recovers the new label contents and sends them to the switch controller. Again, according to the label contents, the switch controller sends control signals to the TLD2 to switch to the correct wavelengths. The new labels L1' and L2' cause the wavelength of the TLD2 to be switched to λ1' (1546nm) and λ2' (1542nm) respectively. TLD2 drives the SOA2 that converts the payload signal onto the desired wavelength by cross-gain modulation. P1 converted to 1546 nm will be routed to the destination port. Figure 97(c) shows P1 on the destination port. P2 converted to 1542 nm will be dropped after AWGR. The switched data payload P1 goes to the data receiver for BER measurements. Here, the switching with new labels L1' and L2' emulates the OLSR3 in Figure 96 (a).

Packet by packet bit-error-rate measurements took place on the P1 at each hop. Figure 98 shows the measured data. Each packet is 600ns long with a 200ns guard time, thus each packet period is 800ns. The bit pattern was $2^{15} - 1$ PRBS truncated into the packets. The three curves in Figure 98 are for the optical baseband back-to-back and the payload signals after one and two hop OLSR, respectively. The signal after one hop shows about 0.7 dB power penalty compared to the baseband payload signal.

However, a negative power penalty about 0.2 dB at BER=1e-9 appears after 2 hop OLSR, which is mainly due to the 2R regeneration in the SOA-based MZI WC and the decrease of the received average power after two packet-droppings. The eye diagrams of the switched payload are shown as the insets in Figure 98.   All eye diagrams show clear openings.   The signal-to-noise ratio was higher for the second hop compared to the first hop due to the 2R regeneration of MZI WC. Also the XGM based SOA wavelength converters invert the logic of the signal which leads to the change of the average power of signal. This results in a higher average power for the first hop. For the second hop the second XGM wavelength converter will invert the logic back to normal. For these reasons, the power received by the date receiver corresponds to different ratios of the real packet power for the 1-hop operations and the 2-hop operations. The combination of the 2R regeneration and the optical power change leads to the negative power penalty for the 2-hop operation.
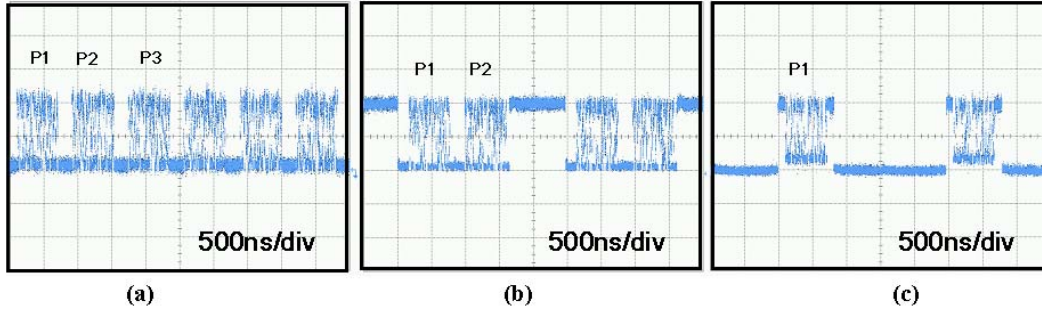

Figure 97:   Scope traces showing (a) the incoming P1, P2 and P3, (b) P1 and P2 sent to the second hop after dropping P3, and (c) P1 at the final destination output after two-hop OLSR.
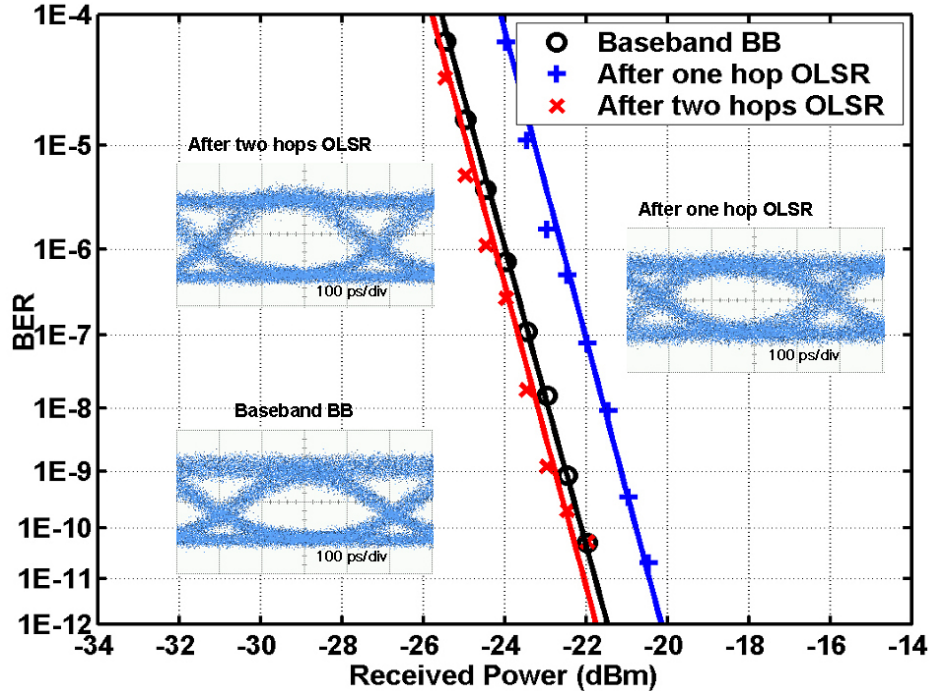

Figure 98:   BER test results of the cascaded OLSR
(Insets: eye diagrams of the baseband payload and signals after OLSR)

5.3.4.2 Summary

We have demonstrated for the first time, to our knowledge, the error-free multi-hop cascaded operation of an all-optical label routing system with optical label swapping. The experiment emulated optical packet switching through 2 hops in the network. Experiment results show regenerating optical label switching with label swapping and 2R packet regeneration. The two-hop routing OLSR system demonstrates negative power penalty of 0.2dB at BER=1e-9 for data packets.

5.3.5 *MZI Wavelength converter/LD/SOA/EA Modulator driver boards*
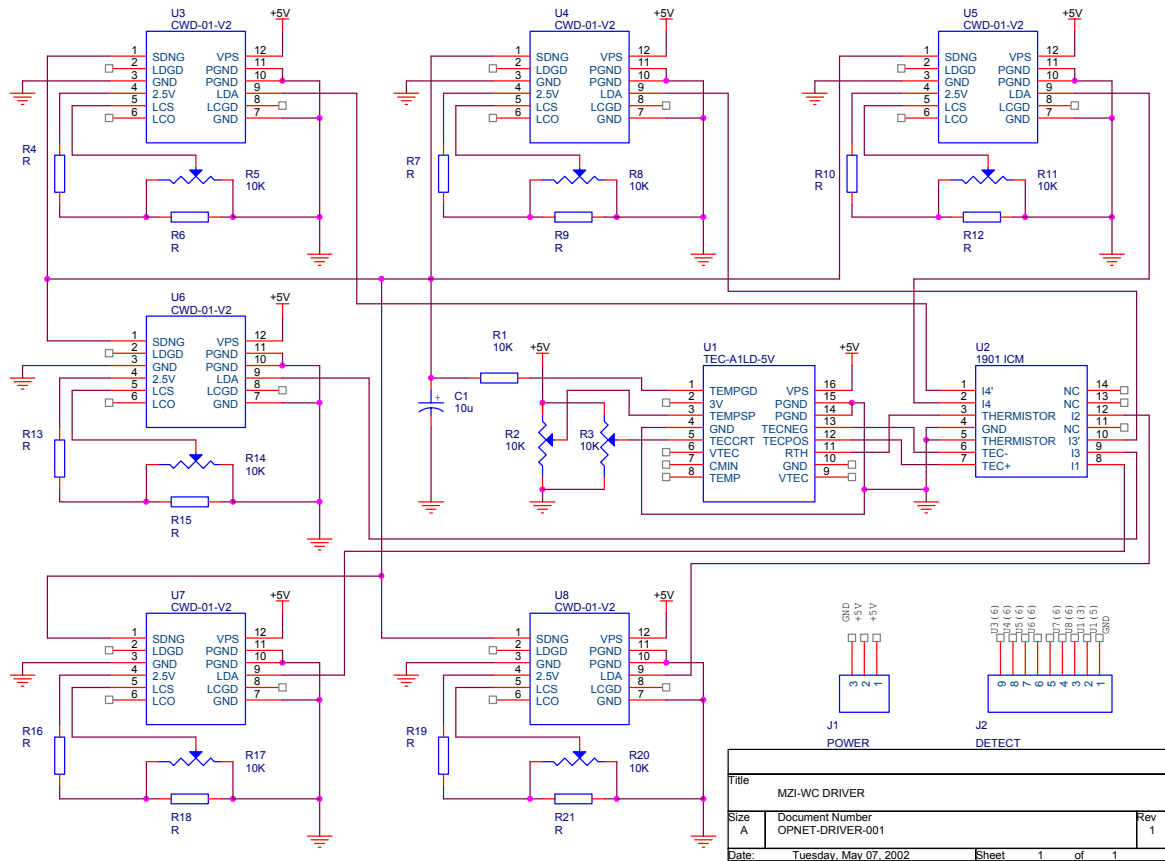#### 5.3.5.1 Schematic and layout for the MZI WC driver

Figure 99:   Schematic graph of MZI wavelength converter driver board

Figure 99 shows the schematic graph of MZI wavelength converter driver board. Current source is Analog Technologies CWD-01-V2 chip with 500mA maximum current. TEC controller is Analog Technologies TEC-A1LD-5V controller chip with 5V maximum TEC voltage. TEC controller will give out shutdown signal to current sources when the temperature stabilization control fails. All the currents and temperature settings will be shown by measure of probes. Temperature setting, TEC voltage control and current settings will be   achieved by adjusting adjustable resistors. Figure 100 is the layout of the PCB for this driver. The future improvement includes display, high frequency circuitry design for modulation of current. This is the final version for fabrication after some minor changes of connection and resistor values.
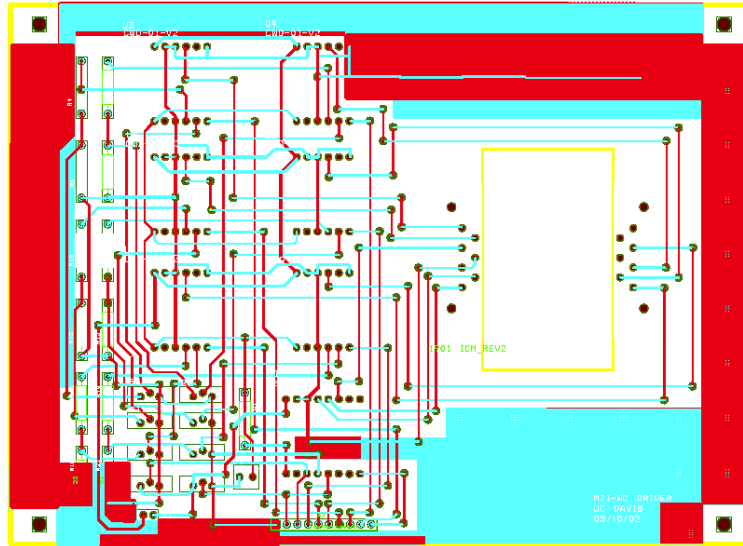
Figure 100:   Layout of PCB of MZI wavelength converter driver

The driver board has been tested and meets all the requirements. MZI wavelength converters based on the driver board and mount are successfully implemented in the Tektronix experiment and show good performance.

5.3.5.2 Schematics and layouts for the LD/SOA/EA Modulator drivers

Figure 101 and 102 show the schematic graph and layout of laser diode driver. Figure 103 and 104 show the schematic graph and layout of SOA driver board. Figure 105 and 106 show the schematics and layout of EA modulator driver board.
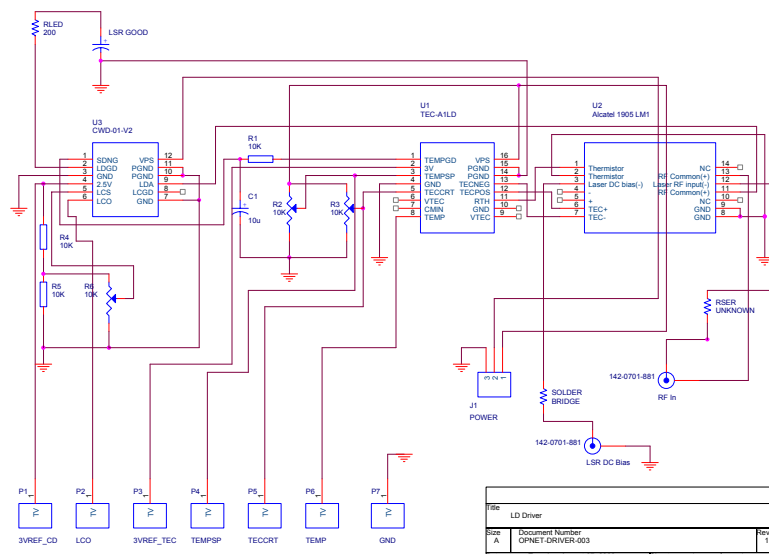

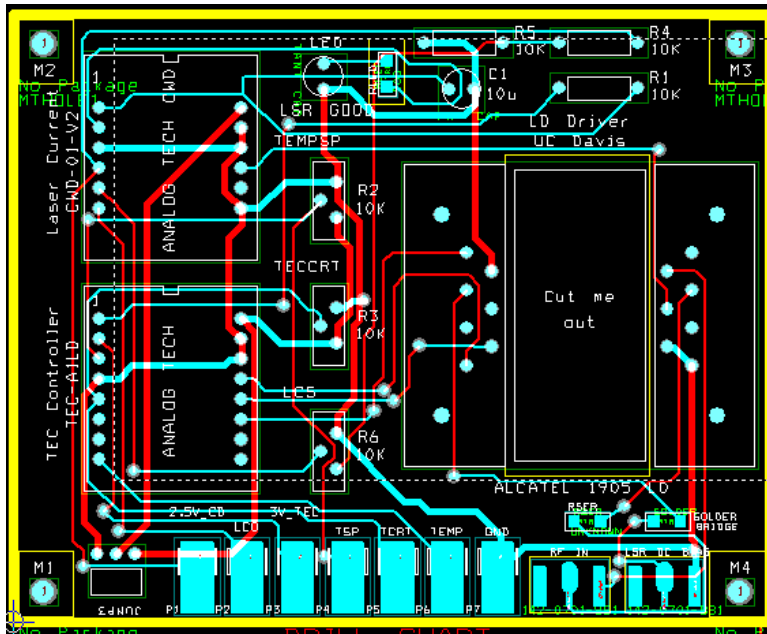Figure 101:   Schematic graph of LD driver board
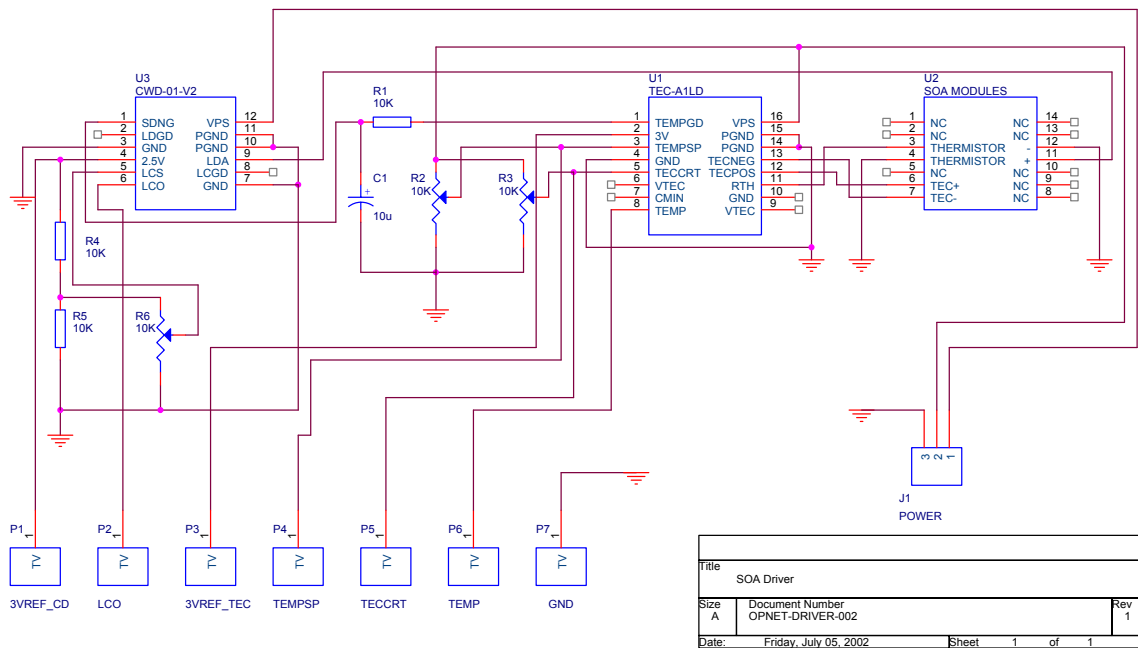
Figure 102: Layout of LD driver board

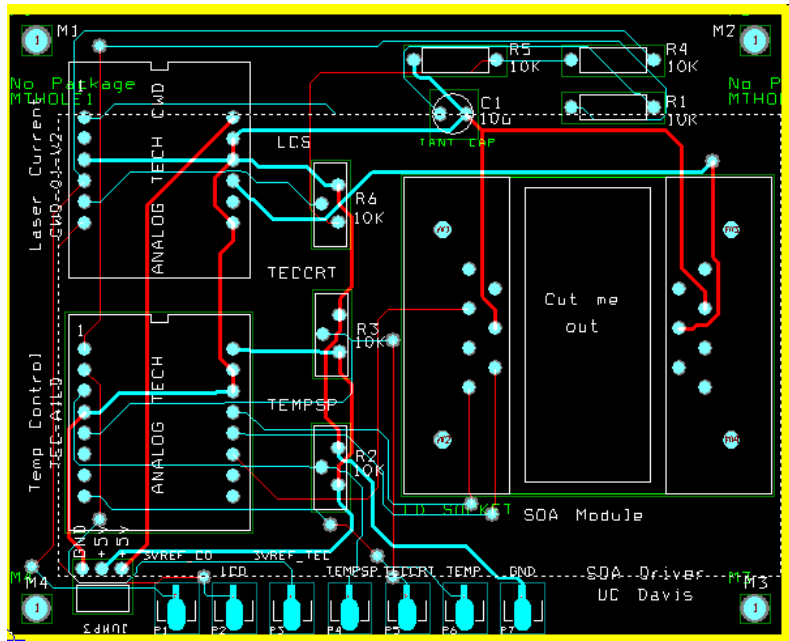Figure 103: Schematics of SOA driver board
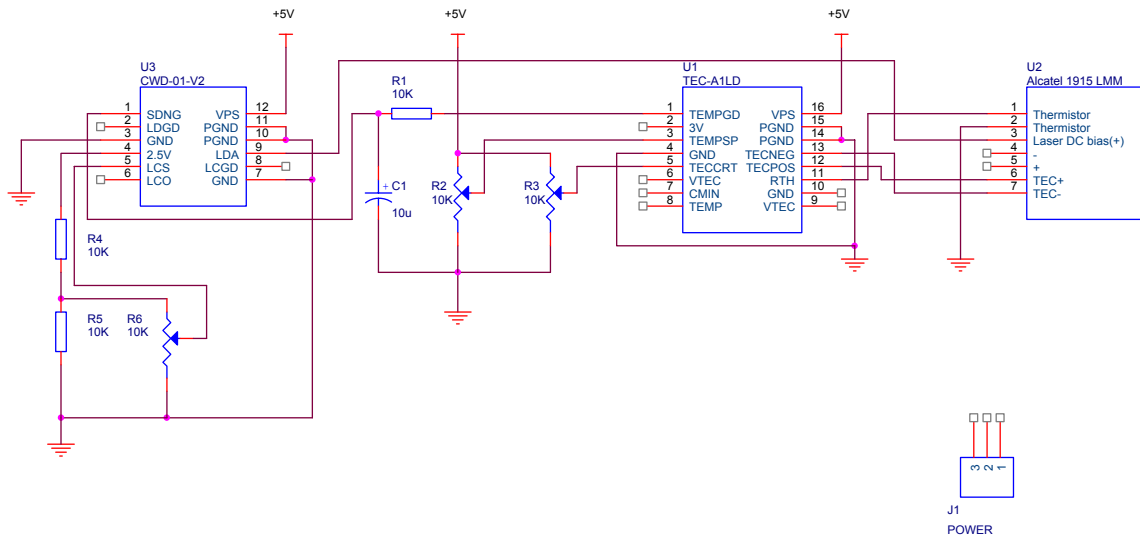
Figure 104:   Layout of SOA driver board

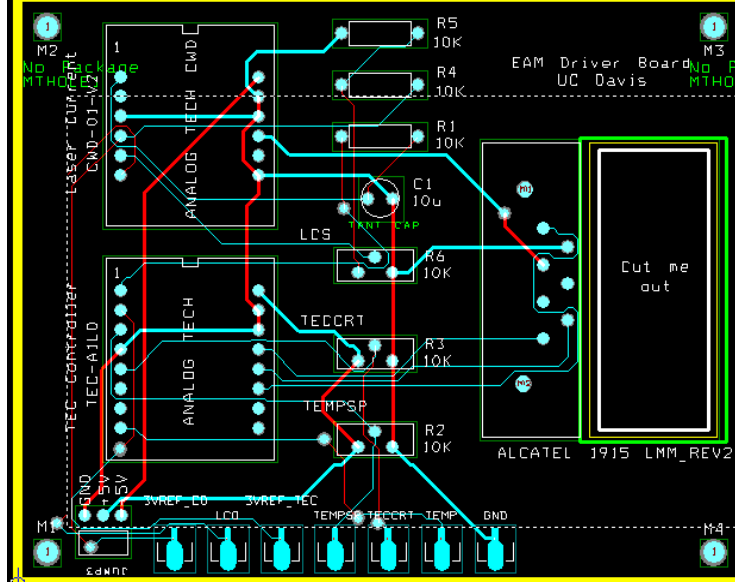Figure 105: Schematics of SOA driver board

Figure 106:   Layout of EA Modulator driver board

All the boards have been tested by using dummy pump laser and work well. They are all ready for integration with device and mount to make compact modules.

5.3.6 Optical gain-clamping EDFA calibration experiment
5.3.6.1 Experiment setup and results

The EDFA calibration experiment is to investigate the gain dynamics of gain-clamping (GC) EDFA, parameters that affects the performance of EDFA and try to find a better solution for gain clamping of EDFA. Figure 107 shows the experimental setup.
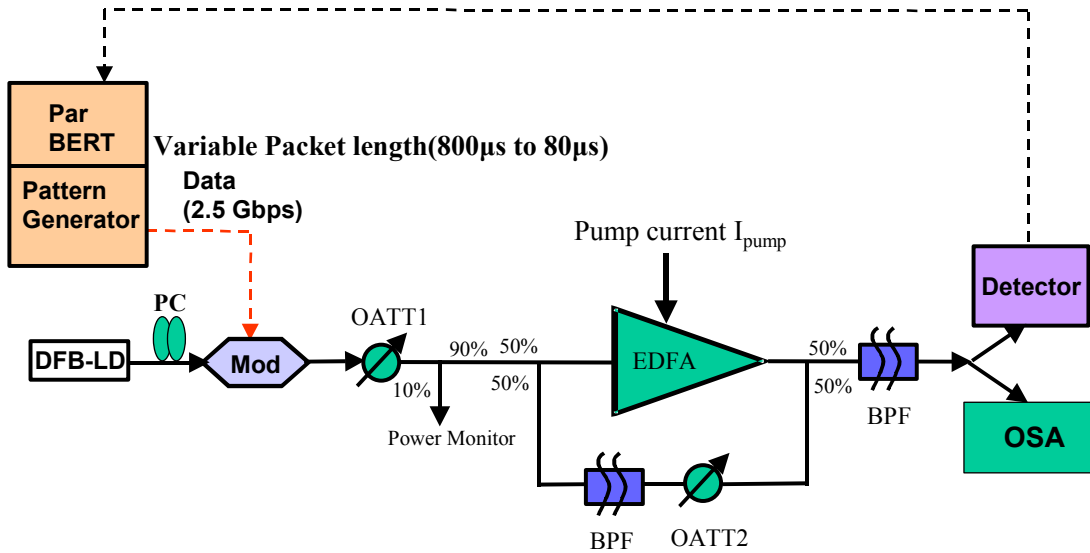

Figure 107:   Experimental setup of gain-clamping EDFA calibration

2.5Gb/s data modulates the LiNbO3 modulator and generates the optical signal. The pattern used for data is the packet with 50% duty cycle and with variable total period lengths from 800μs to 80μs. The bit pattern is 2^{15}-1 PRBS truncated into the packet.   The structure of EDFA is as shown in Figure 108 which is a dual pump structure. 50% of power from output of EDFA is fed into the feedback loop. The feedback signal will pass optical attenuator and tunable bandpass filter and go into EDFA again through another 50:50 coupler. This is an optical feedback structure to achieve gain clamping. The output signal is sent to OSA and BERT for optical signal analysis and bit error rate testing. In the experiment, the parameters that can be adjusted includes pump current, input signal power (by changing OATT1), feedback ratio (by changing OATT2), lasing wavelength (by changing center wavelength of BPF) and packet length. The lasing wavelength is set to 1533nm because it's not a critical parameter. The performance of gain-clamping changed little by changing the lasing wavelength from 1530nm to 1540nm. The wavelength of the input signal is 1550nm. The pattern lengths used are 800μs and 80μs. To simplify the experiment co-pumping laser is turned down to make EDFA single pump amplifier.

Dual-pump EDFA structure
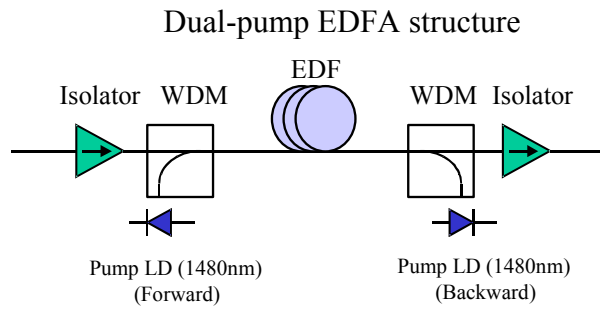
Length of EDF: ~20m

Figure 108:   Structure of EDFA
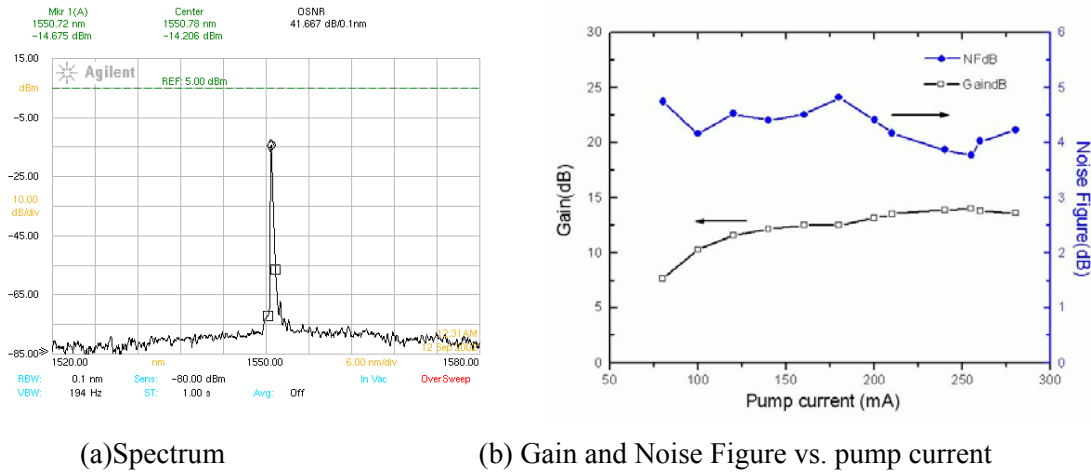
(a)Spectrum   (b) Gain and Noise Figure vs. pump current
Figure 109:   Output spectrum and gain & NF performance

Figure 109   shows the spectrum and gain & noise figure (NF) performance while changing pump current. From the curve we can conclude that pump current about 250mA can maximize the gain and minimize the noise figure. So this pump current is fixed for the future measurement.

Figure 110 & 111 show the spectrum and change of the gain vs. input optical power. From the figures we can conclude that saturation power of GC-EDFA is about –6dBm which can maximize the gain and

stabilize it. From the spectrum we can see clearly when the input signal is weak (lower than -15dBm) the lasing wavelength will dominate; the signal wavelength will dominate while the input signal power is higher than several mW. Both wavelengths will both survive while signal power is between these two values. From the later experiment we can also see that the feedback ratio will change this range. From Figure 111 we can see that the saturation power will decrease from 0dBm to –8dBm while the feedback attenuation (Att2) changes from 0.7dB to 6.5dB (feedback ration from 21% to 5.5%).

(a) P_{in}=-15dBm   (b) P_{in} = -5dBm   (c) P_{in} = 4.7dBm
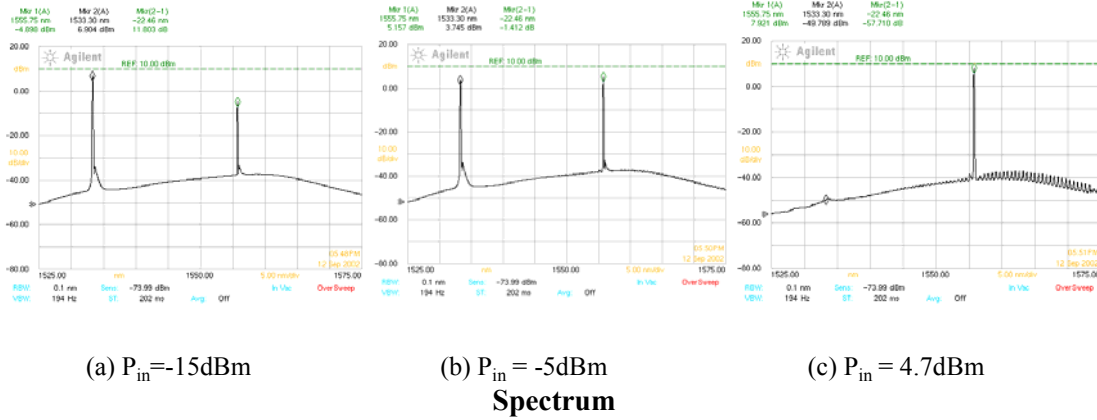
Spectrum

Figure 110:   Spectra of EDFA output for different input power
(Saturation power of GC-EDFA is about –6dBm)
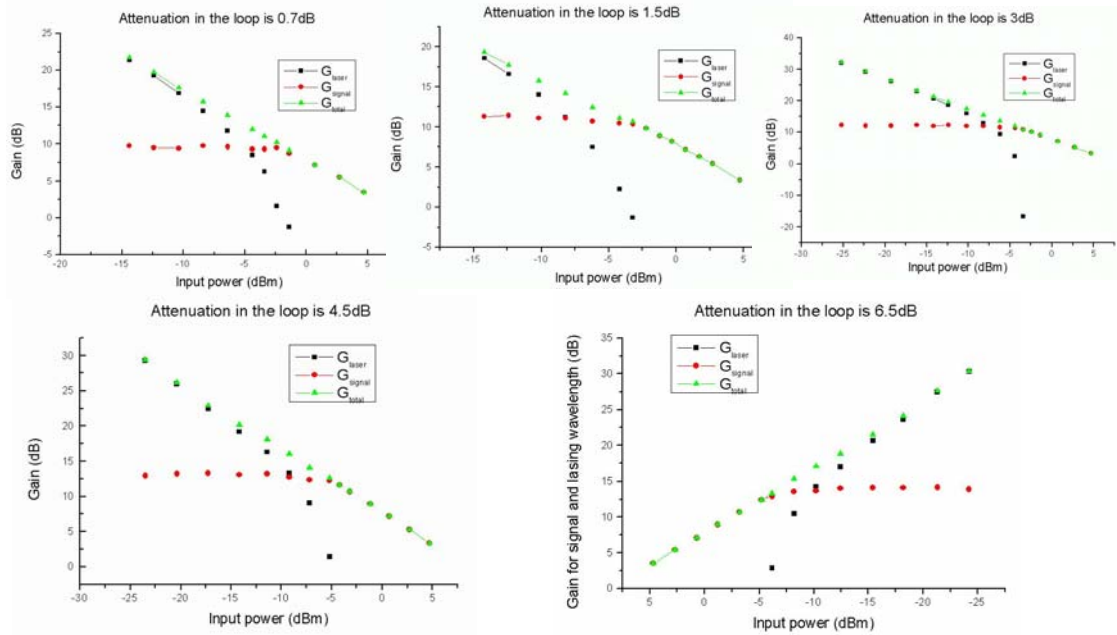

Figure 111:   Gain vs. input power

Figure 112 & 113 show how feedback ratio affects the gain and NF.   Figure 113 is the combination of two groups of curves in Figure 112.   From the figures we can conclude that at 6-7% feedback intensity level noise figure reaches minimum and gain is also relatively high at this level. So this feedback ratio is about the optimum point.

Figure 112:   Gain & NF vs. feedback ratio
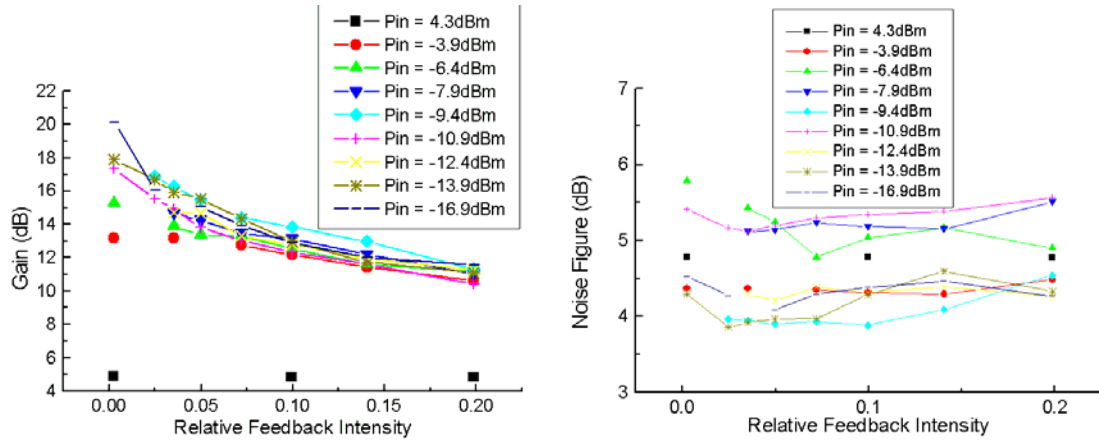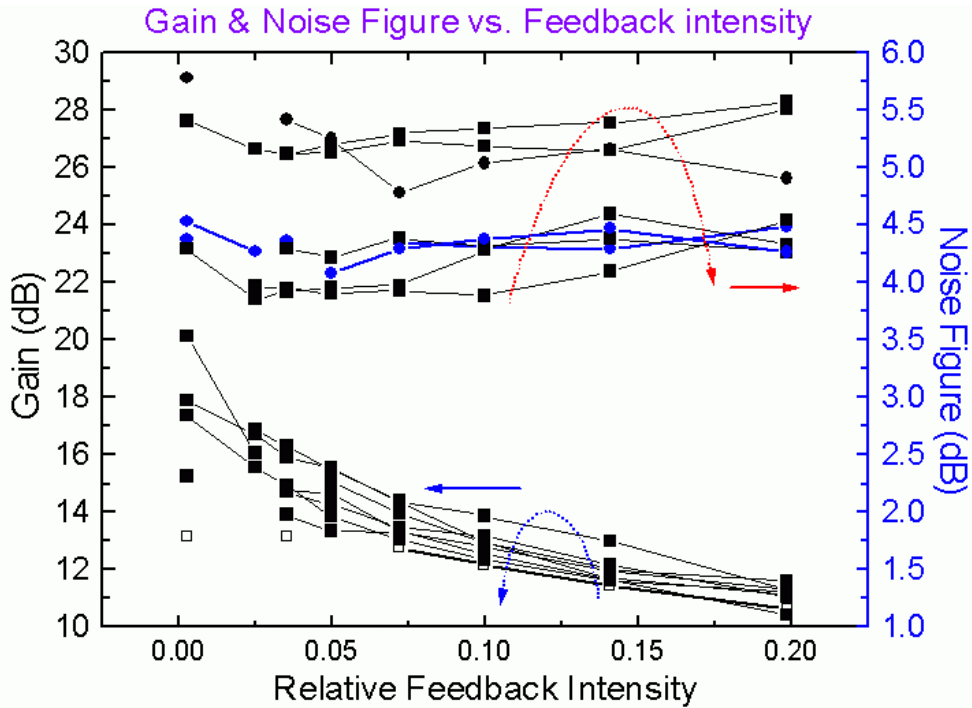

Figure 113:   Combination of gain & NF vs. feedback ratio

Below are the measurements for the dynamic performance of this gain-clamping EDFA. Figure 114 -118 show the waveforms for different feedback ratio for input power at –3.4, -6, -9, -14 and -18dBm respectively. Figure 119 - 122 show the packet by packet bit error rate testing results for input power=-3.4, -6, -9, -14dBm.

Figure 114: Waveforms for input power at –3.4dBm (Att is the attenuation in the ring laser loop which represent the different feedback level. The insets show the clear transient by using only long "1" and long "0" instead of PRBS data.)

Att = 0.7dB   Att = 6dB   Att = 13dB
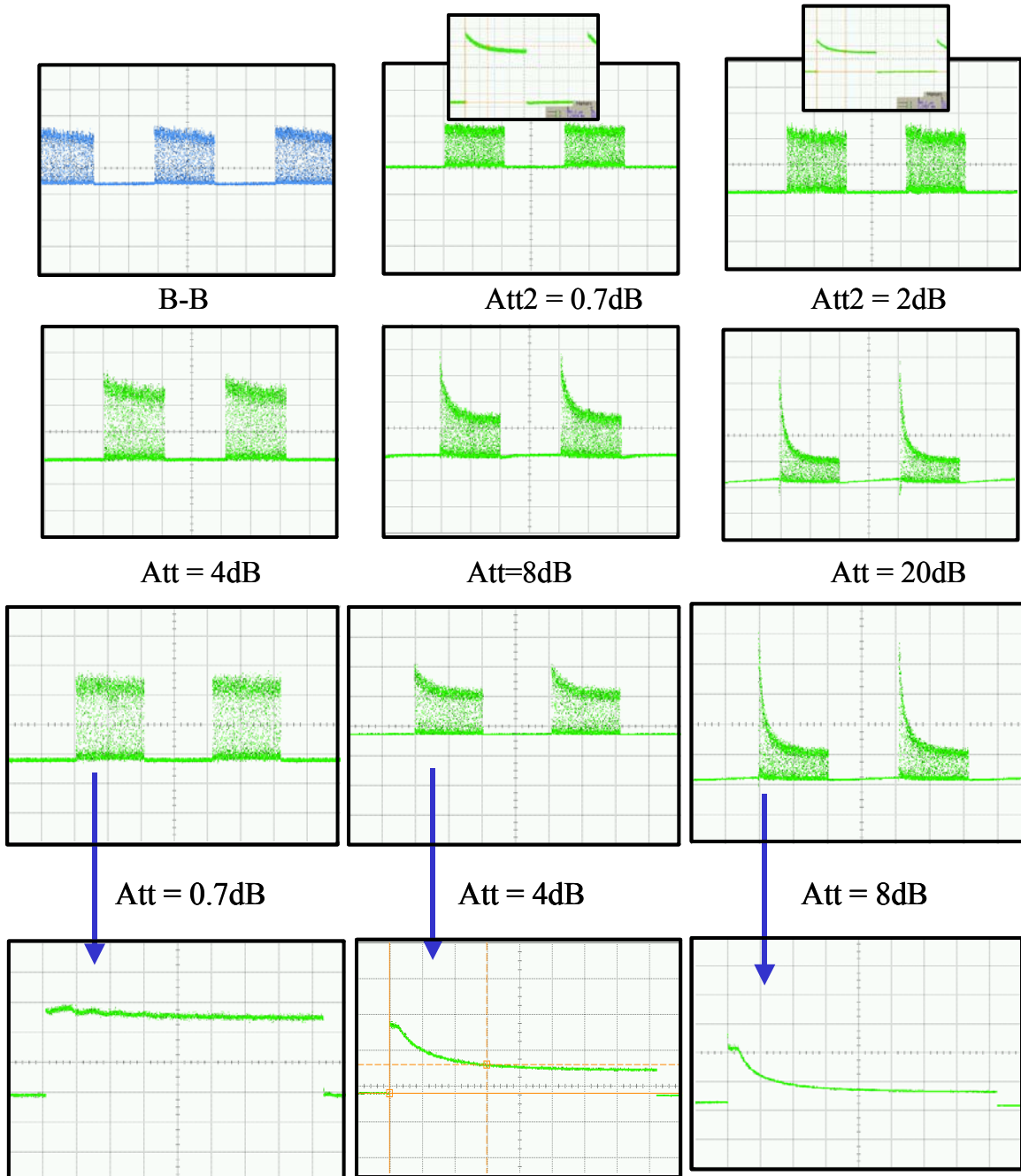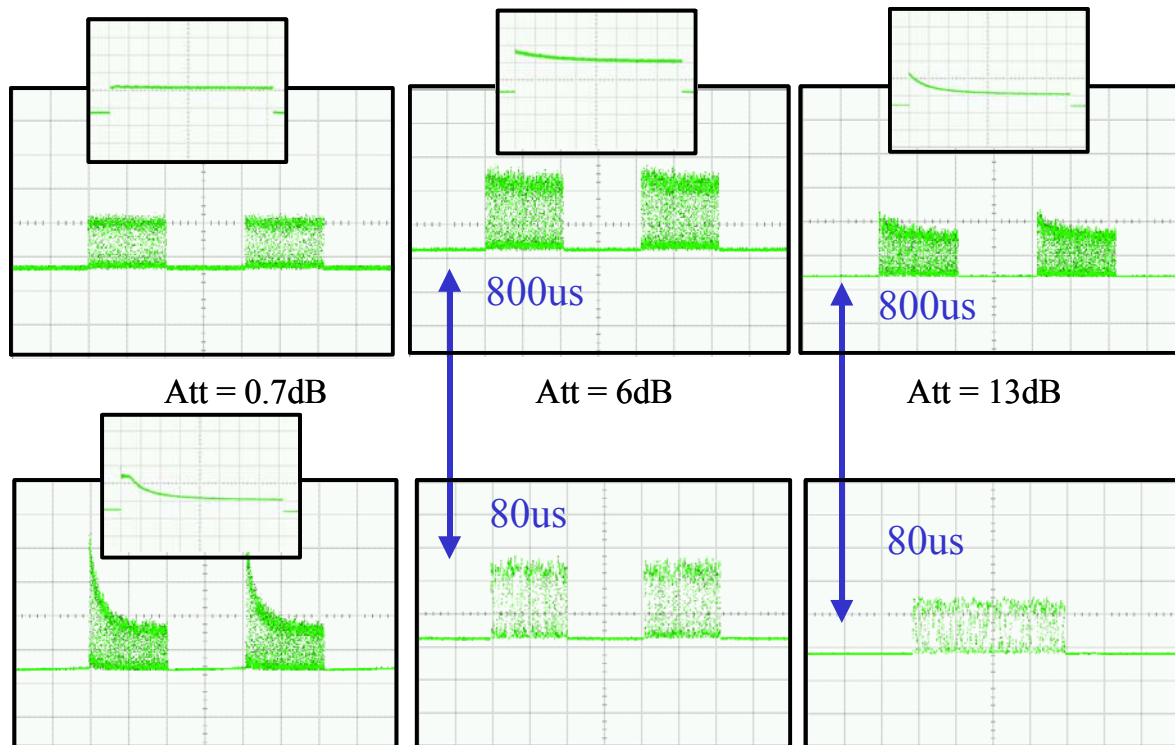
Figure 115: Waveforms for input power at –6dBm (Att is the attenuation in the ring laser loop which represent the different feedback level. The figures at the bottom show the clear transient by using only long "1" and long "0" instead of PRBS data)

Figure 116: Waveforms for input power at –9dBm (Att is the attenuation in the ring laser loop which represent the different feedback level. The insets show the clear transient by using only long "1" and long "0" instead of PRBS data.)

No GC   Att = 2dB   Att = 10dB
~200us   ~10us   ~50us

Figure 117: Waveforms for input power at –14dBm (Att is the attenuation in the ring laser loop which represents the different feedback level. The insets show the clear transient by using only long "1" and long "0" instead of PRBS data.)
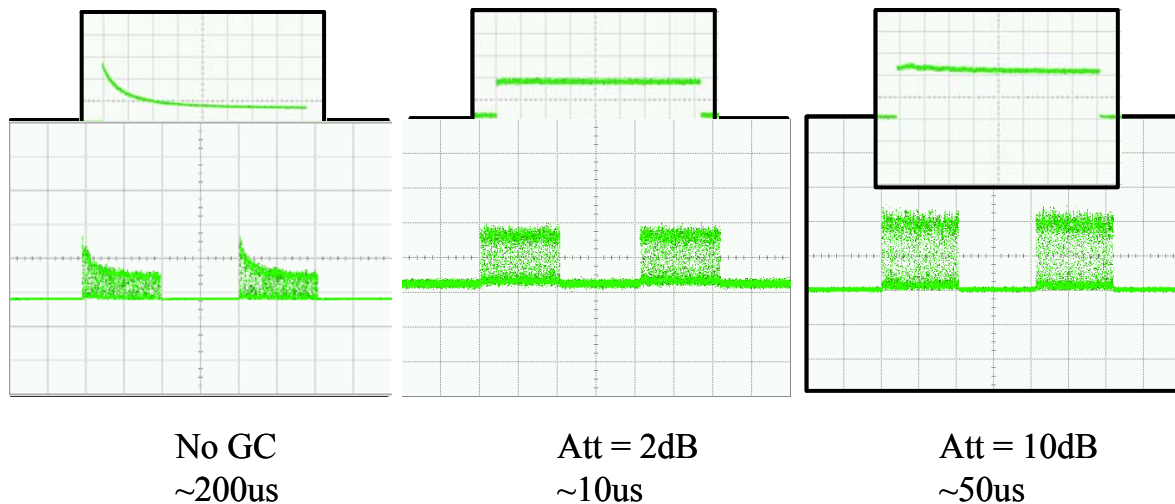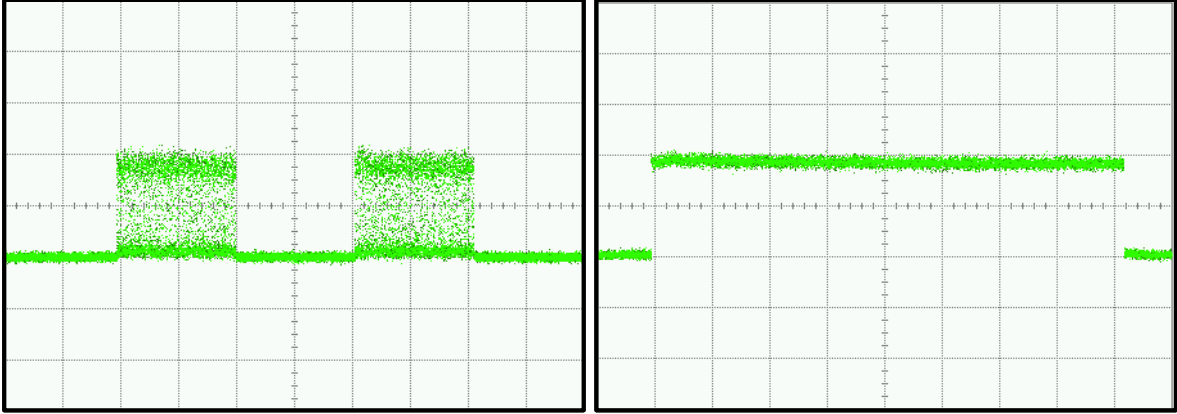
92

Figure 118: Waveforms for input power at –18dBm (Att=10dB, transient is about 20μs, Rx sensitivity(Error free)=-28.3dBm. The right figure shows the clear transient by using only long "1" and long "0" instead of PRBS data.)
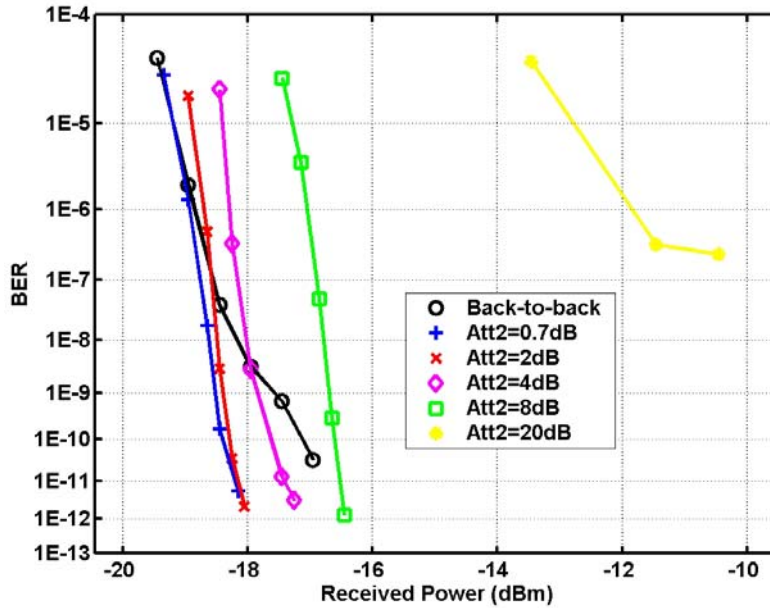

Figure 119: BER curve for input power at –3.4dBm

From waveforms, BER curves and static results before we can get the following conclusions:
1. Packet length affects the gain dynamics of EDFA strongly. If the packet length is comparable to EDFA life time (~ms), a strong gain transient will degrade the signal and result in more bit errors. For the 800μs long packet, error floor at $1*10^{-5}$ or $1*10^{-6}$ exists for the cases that input power is higher than –6dBm. For 80μs long packet, no error floor shows up even the input power is high.

2. For long packet, stronger feedback is necessary for higher input power to achieve a fast transient and a better performance. For instance, if input power is –6dBm a feedback ratio of 20% (Att2=0.7dB) results in the best BER performance as shown in Figure 120.   But if input power is –14dBm, a feedback ratio pf 2.5% (Att2=10dB) will lead to a better result than stronger feedback as shown in Figure 122.

3. Comparing the waveforms and BER results, we can conclude that the input power should be between – 6 to –14dBm to get a good performance if we set the feedback ratio at 12% (Att2=3dB); If the input

power is low, a weaker feedback is necessary to get enough gain for signal and better performance for gain dynamics.

4. Optical gain-clamping method is good in sense of transient suppression. The transient time can be suppressed to less than 10μs (for input power=-14dBm and Att2=2dB). But the dynamic range for input power is small (less than 10dB).
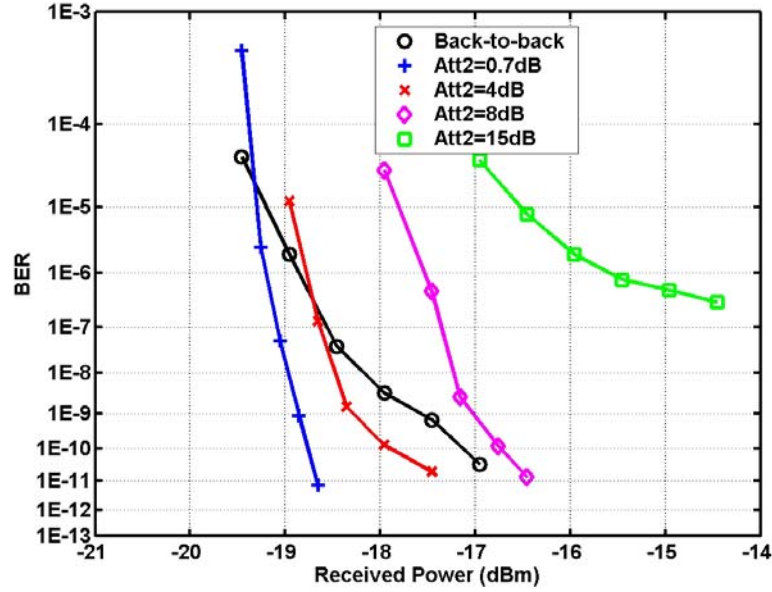
Figure 120:   BER curve for input power at -6dBm
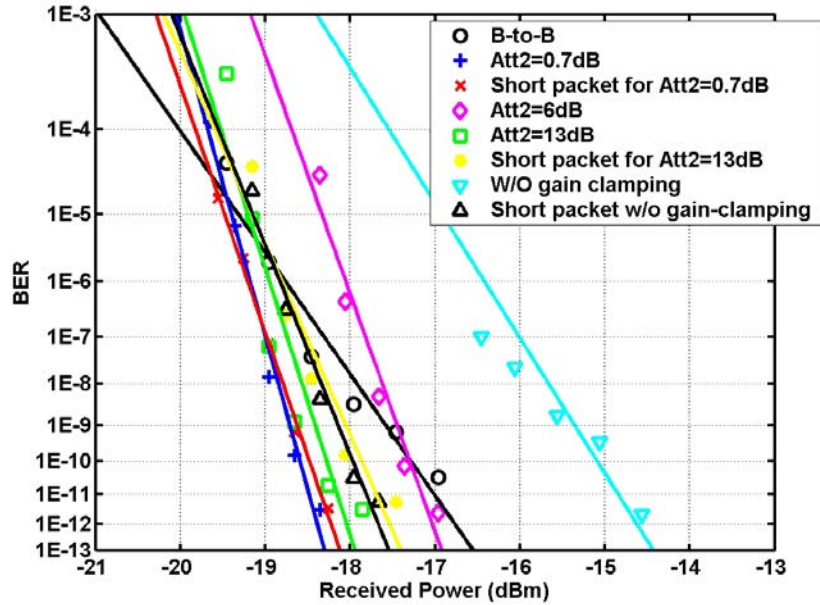
Figure 121:   BER curve for input power at –9dBm
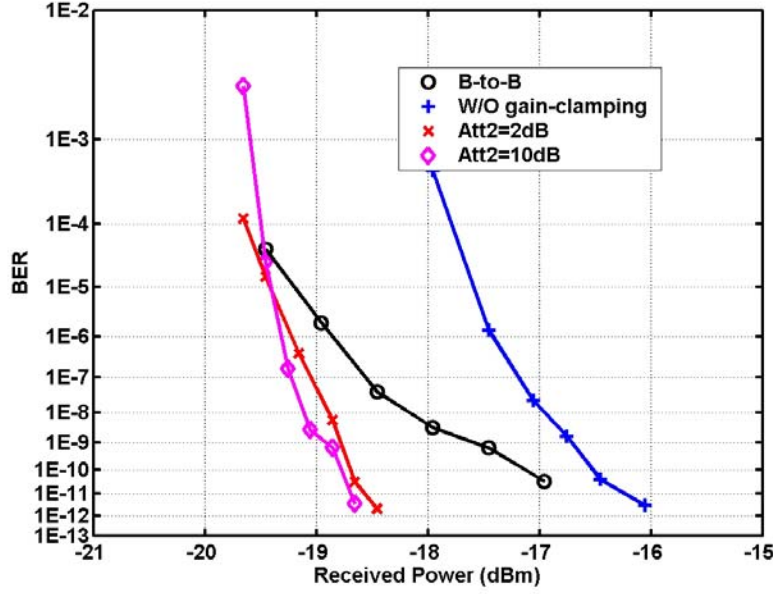
Figure 122:   BER curve for input power at –14dBm

5.4. Device Design and Characterization
5.4.1 Client Interface
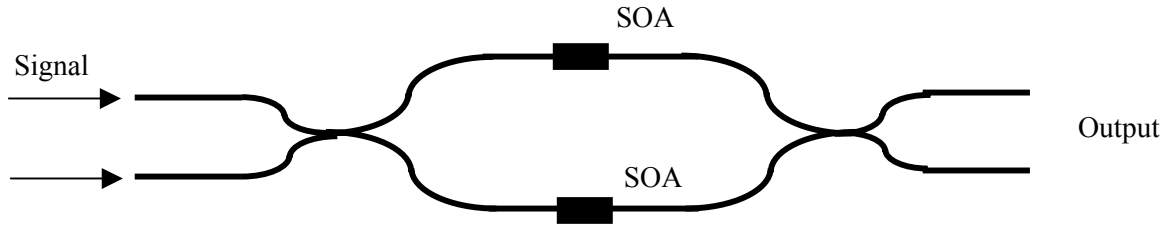
Figure 123:   Schematic diagram of the MZI device

The test of the MZI device was reported.   Figure 123 shows the structure of the MZI device. It consists of two SOA at each arm. The total loss (including coupling loss and waveguide loss) is about 20dB for the testing waveguide on the same sample. The total loss for the MZI with no current injection is about 30dB.   And the total optical power for the ASE of one SOA in the MZI structure is about –34dBm. Figure 124 shows the spectrum of the ASE.
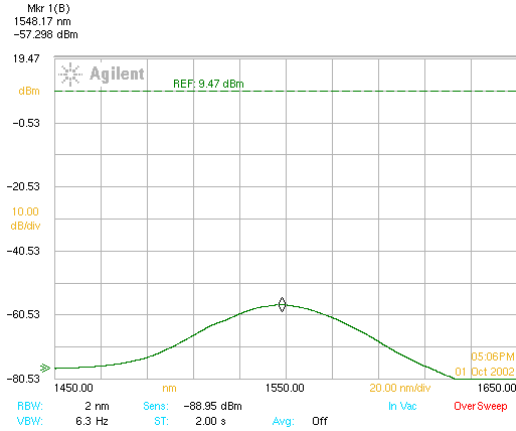
Figure 124:   Spectrum of ASE

And we also tried to input different wavelength light into the MZI structure. Figure 125   show the spectra of the outputs (the injected current is 100mA, the input light wavelengths are 1550nm and 1560nm respectively):
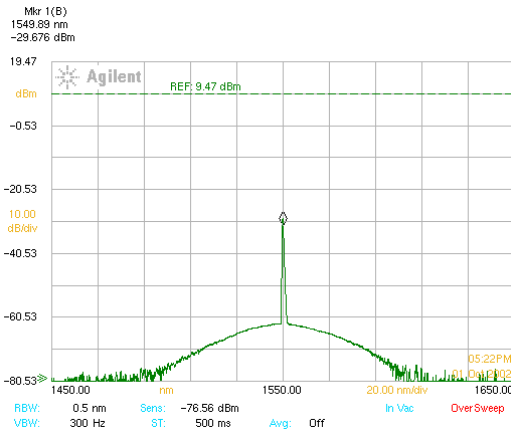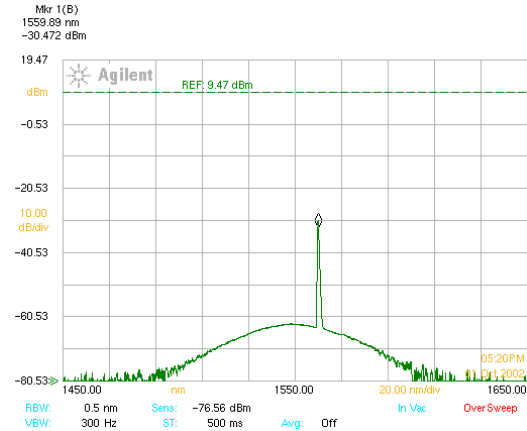
(a)   (b)

Figure 125:   Optical spectra of the outputs (a)input wavelength 1550 nm (b) 1560 nm

From the output, We weren't able to see any amplify effect (stimulated emission) from the SOA. It needed to be invested in the future.

5.4.2 Simulation of the OTDM system

We have correct one error in the old program, which shows better performance of the Mach-Zehnder pulse detector. Figure 126 show the output signal pulse for the matched code and un-matched code.
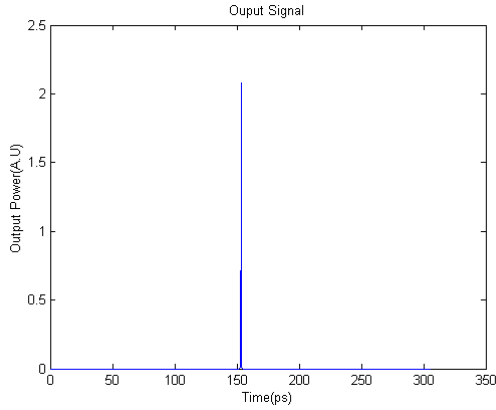
(a) Total energy = 0.9992   (b) Total energy = 0.0542
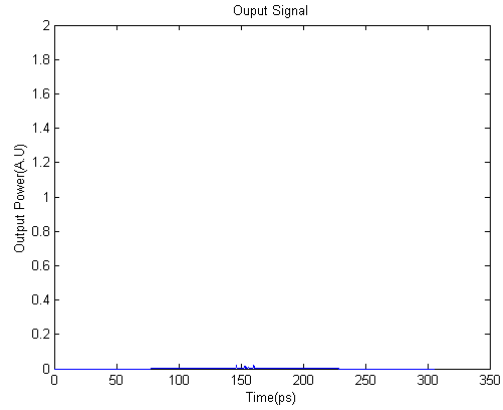
Figure 126:   Outputs signal pulses for the matched and unmatched code.

We also start to apply multi-channel interference and test the BER. The electrical field from multi-channel interference can be seen in the Figure 127:
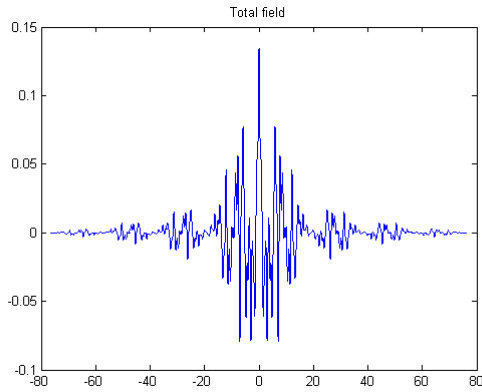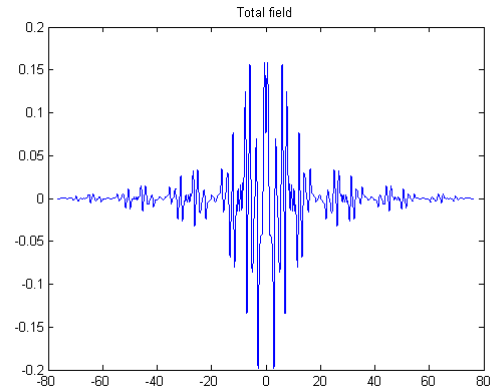
(a) 4 channels   (b) 8 channels
Figure 127:   Electrical fields from multi-channel interference

Figure 128 show the outputs signal for the case of four channels and eight channels interference respectively.
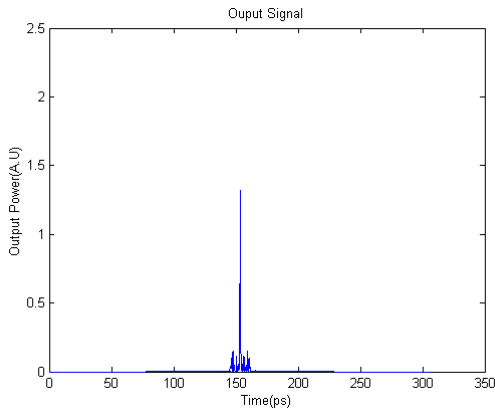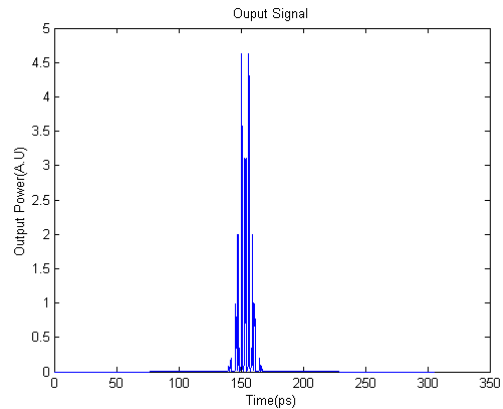
(a) 4 channels   (b) 8 channels
Figure 128:   Outputs signal for the case of four channels and eight channels interference

97

The BER test result for 127 bit PRBS is shown Figure 129. Y-axis shows the error count and X-axis shows the number of channels.
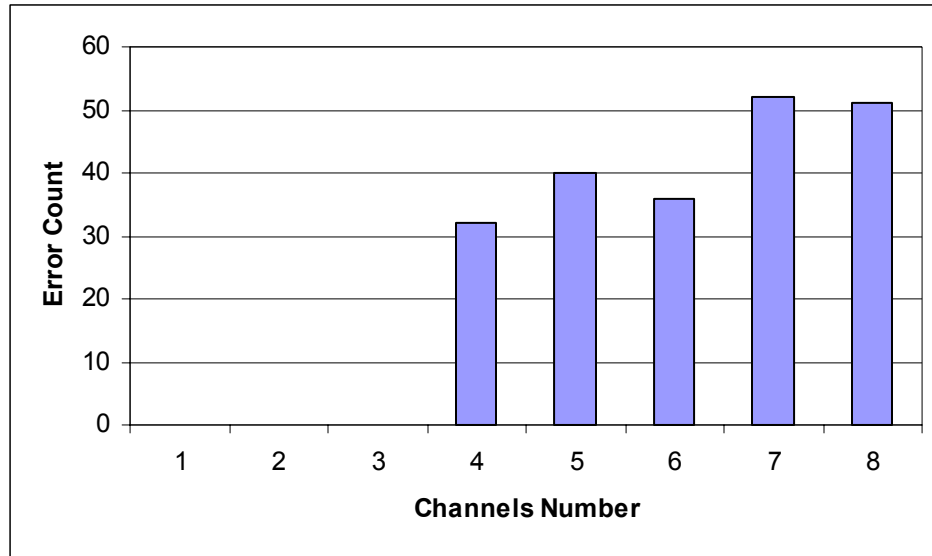
Figure 129:   BER test result for 127 bit PRBS

The error rate is high. BER for longer PRBS is being tested right now. New coding scheme may need to be utilized in order to solve the problem of strong interference form undesired channel.

5.4.3 Wafer bonding

The wafer loading is a critical step in the bonding procedure. It is the only step whose had not been finalized. Since this is the only step, during which the chamber is opened to the outside atmosphere. Without careful arrangement, both the wafer surfaces and chamber environment can be contaminated. Ideally, we wanted the wafers to be placed together with bonding surfaces closely facing each other after preliminary cleaning, to avoid direct contact with the atmosphere. The two wafers would also be aligned at this point. At the same time, the chamber is purged with nitrogen. After wafer loading and chamber closing, the two wafers can be separated by vacuum on both the top and bottom holding plates. The above wafer loading method was tested, and two other loading methods were proposed according to test results.

In the current chamber design, there is no vacuum on the bottom holding plate. Before chamber modification for adopting the above loading procedure, some tests were performed. An aluminum plate with vacuum holes was made as an adaptor placed on the bottom plate (Figure 130).   After RCA cleaning, the two wafers were placed face-to-face on the adaptor and aligned to each other. The vacuum on the adaptor was used to hold the aligned wafers. To make the two wafers stay together, slight pressing was applied so that the wafers were bonded by Van der Waals force. Then the wafers were loaded into the chamber along with the adaptor. The wafers were tried to be separated inside the chamber by the vacuum on the top plate and the adaptor. Because of the vacuum line connected to the adaptor, the chamber could not be closed at this point. Since the chamber was constantly purged by nitrogen, the possibility of wafer-surface-contamination by the atmosphere could be greatly reduced.

However, it was difficult to separate two wafers with good surface quality, because of the strong Van der Waals bonding. At the beginning, used wafers were used for testing and the proposed procedure was always successfully accomplished. With new wafers, the Van der Waals bonding was too strong to be broken by the vacuum.

Figure 130:   Schematics of wafer separation after loading

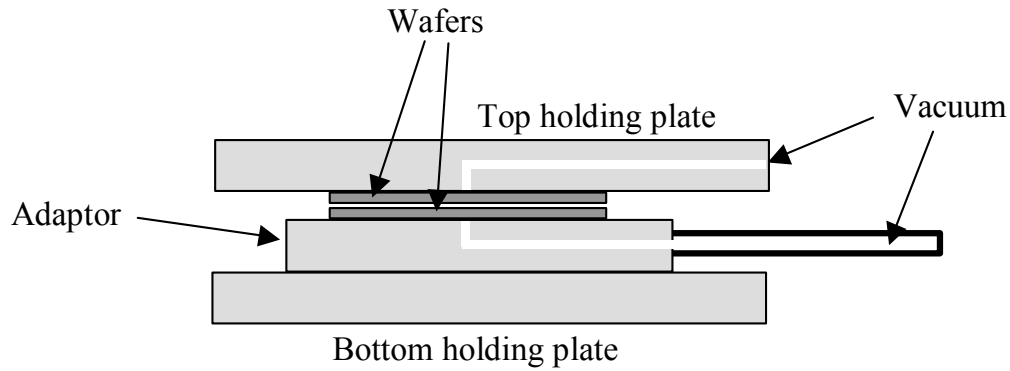One possible solution is to heat the wafer during separation, because the Van del Waals force is weaker at high temperatures. The wafer-bonding chamber is equipped with IR heater; therefore, this method can be easily accomplished. However, due to pump breakdown, it has not been tested.
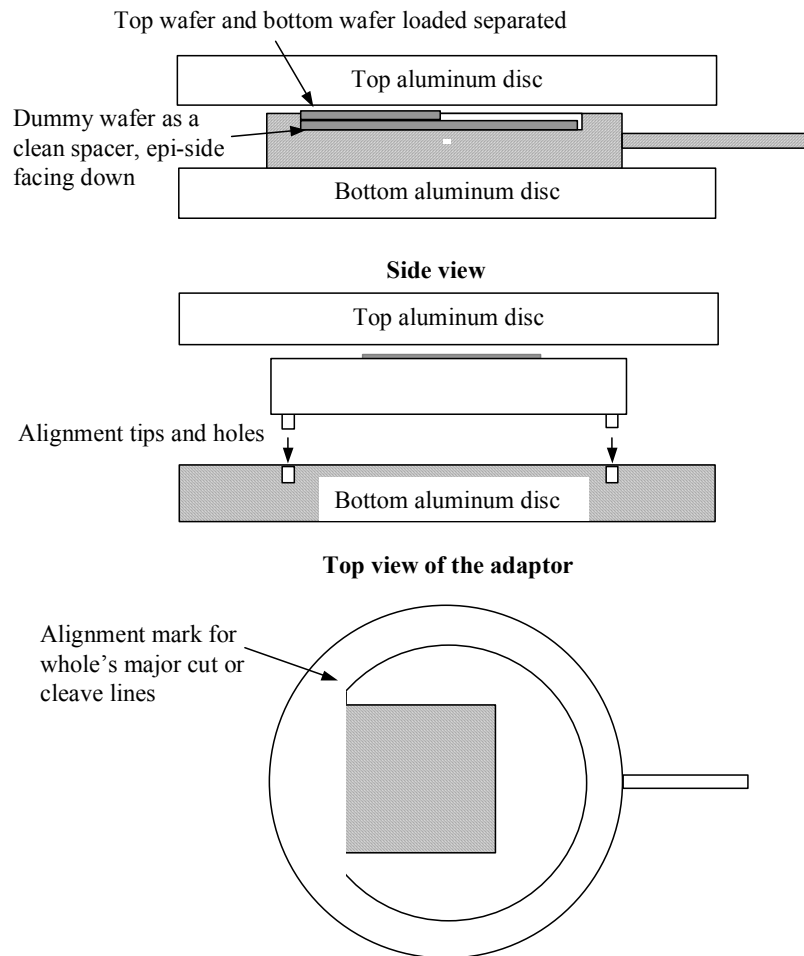

Figure 131:   Schematics of the proposed tool for loading wafer separately

Another alternative method is to load the two wafers separately. A special tool is proposed and is illustrated in Figure 131.   In this method, the top wafer is loaded first, and then the bottom wafer.   The loading tool will be left in the chamber during bonding processes. Both wafers will be aligned with the mark on the tool. With the help of alignment tips and holes on the tool and bottom plates, respectively, the position of the tool on the bottom plates will be same each time it was loaded. In this way, the two wafers can be aligned even they are loaded separately. A dummy wafer, which will be RCA cleaned each time before loading, will be used to avoid direct contact with the loading tool.

6. Publications and Presentations

PUBLICATIONS:

1. Shun Yao, Biswanath Mukherjee, S. J. Ben Yoo, Sudhir Dixit, "A Unified Study of Contention-Resolution Schemes in Optical Packet-Switched Networks," submitted to *IEEE/OSA Journal of Lightwave Technology* (2001).
2. Jing Zhang, Keyao Zhu, Laxman Sahasrabuddhe, S. J. Ben Yoo, and Biswanath Mukherjee, "On the Study of Routing and Wavelength Assignment Approaches for Survivable Wavelength-Routed WDM Mesh Networks," accepted for publcation *Optical Networking Magazine*, Kluwer Academic Publishers (2002).
3. Shun Yao, Fei Xue, Biswanath Mukherjee, S. J. Ben Yoo , Sudhir Dixit, "Electrical Ingress Buffering and Traffic Aggregation for Optical Packet Switching and Its Effect on TCP Performance," accepted for publication *IEEE Communications Magazine* (2002)
4. S. J. B. Yoo, Hyuek Jae Lee, Yanda Zhang, S. K. H. Fong, Vincent Hernandez, Vincent K. Tsui, Srikanth Vaidianathan, Katsunari Okamoto, Shin Kamei, "Rapidly Switching All-Optical Packet Routing System with Optical-Label Switching and Optical-Label Swapping using Rapidly Tunable Wavelength Conversion and a Uniform-Loss Cyclic Frequency AWGR," *to be published in IEEE photonics technology letters* (2002).
5. S. J. B. Yoo, "Rapidly Switching Optical Packet Routing System with Wavelength, Time, and Space Domain Contention Resolution**,"** Technical Digest, *Photonics in Switching Conference 2002,* Cheju, Korea (2002) (**Invited).**
6. Jing Cao, Zhong Pan, Minyong Jeon, Yash Bansal, Julie Taylor, Venkatesh Akella, Shin Kamei, Katsunari Okamoto, S. J. B. Yoo, "Cascaded Operation of an Optical Packet Routing System with Optical-Label Switching and 2R-regeneration," accepted for publication in *European Conference on Optical Communication* , Kopenhagen, Denmark (2002).
7. S. J. B. Yoo, "All Optical Label Switching: Architecture, Protocol, and System Demonstration", to be published in Technical Digest, 7th OptoElectronics and Communications Conference, Yokohama, Japan (2002) (**Invited).**
8. S. J. B. Yoo, "Optical-label switching, MPLS, MPLambdaS, and GMPLS," *to be published in Special Issue on: Prospects and Challenges: Next Generation Optical Network Architectures, Optical Networking Magazine*, Kluwer Academic Publishers (2002) (**invited).**
9. Shun Yao, Biswanath Mukherjee, S. J. Ben Yoo, Sudhir Dixit, "A Unified Study of Contention-Resolution Schemes in Optical Packet-Switched Networks," submitted to *IEEE/OSA Journal of Lightwave Technology* (2001).
10. Shun Yao, Fei Xue, Biswanath Mukherjee, S. J. Ben Yoo , Sudhir Dixit, "Electrical Ingress Buffering and Traffic Aggregation for Optical Packet Switching and Its Effect on TCP Performance," submitted to *IEEE Communications Magazine* (2002)
11. Shun Yao , Biswanath Mukherjee , S. J. Ben Yoo, and Sudhir Dixit, 'All-optical Packet-switched Networks: A Study of Contention-resolution Schemes,' submitted to IEEE JSAC/APOC

12. Fei XUE, S. J. Ben Yoo, "Self-similar Traffic Shaping at the Edge Router in Optical Packet-switched Networks," *The IEEE International Conference on Communications 2002* (2002).

13. Wushao Wen, Biswanath Mukherjee, and S. J. B. Yoo, "Quality of Service Based Protection in MPLS Control WDM Mesh Networks," *accepted for publication in* Photonic Network Communications, (2001).

14. S. J. B. Yoo, Yash Bansal, Zhong Pan, Jing Cao, Vincent K. Tsui, Steven K. H. Fong, Yanda Zhang, Julie Taylor, Hyuek Jae Lee, Minyong Jeon, Venkatesh Akella, Katsunari Okamoto, Shin Kamei, "Optical-Label Switching based Packet Routing System with Contention Resolution Capabilities in Wavelength, Time, and Space Domains," *in Technical Digest, IEEE/OSA Optical Fiber Communication Conference*, paper #WO2, Anaheim, California, (2002).

15. Fei Xue, Shun Yao, Biswanath Mukherjee, and S. J. B. Yoo, "The performance improvement in optical packet-switched networks by traffic shaping of self-similar traffic," *in Technical Digest, IEEE/OSA Optical Fiber Communication Conference*, paper #WG4, Anaheim, California, (2002).

16. V.J. Hernandez, Z. Pan, J. Cao, V.K. Tsui, Y. Bansal, S.K.H. Fong, Y. Zhang, M.Y. Jeon, S.J.B. Yoo, Brian Bodtker, Steven Bond, William J. Lennon, Hank Higashi, Bryan Lyles, Robert McDonald, "First Field Trial of Optical Label Switching and Packet Dropping on a 477km NTON/Sprint Link," *in Technical Digest, IEEE/OSA Optical Fiber Communication Conference*, paper #TuY4, Anaheim, California, (2002).

17. Zhong Pan, Jing Cao, Yash Bansal, Vincent K. Tsui, Steven K. H. Fong, Yanda Zhang, Julie Taylor, H. J. Lee, Minyong Jeon, Venkatesh Akella, S. J. B. Yoo, Katsunari Okamoto, Shin Kamei, "All-optical programmable time-slot-interchanger using optical-label switching with tunable wavelength conversion and *N* by *N* arrayed waveguide grating routers," *in Technical Digest, IEEE/OSA Optical Fiber Communication Conference*, paper #WM6, Anaheim, California, (2002).

18. S. J. B. Yoo, "Demonstration of optical packet switching using optical-label switching technologies," SPIE Asia-Pacific Optical and Wireless Communications Paper 4582-21 (2001) (**invited).**

19. S. J. B. Yoo, "Optical-label switching, MPLS, MPLambdaS, and GMPLS," SPIE Asia-Pacific Optical and Wireless Communications Paper 4585-02 (2001) (**invited).**

20. S. Yao, S. J. B. Yoo, B. Mukherjee, S. Dixit, "Hybrid contention resolution for an optical packet-switched network with self-similar IP traffic," SPIE Asia-Pacific Optical and Wireless Communications Paper 4585-04 (2001).

21. W. Wen, B. Mukherjee, S. J. B. Yoo, "QoS based protection in MPLS controlled WDM mesh networks**,"** SPIE Asia-Pacific Optical and Wireless Communications Paper 4585-17 (2001).

22. Hyuek Lee, Vincent Hernandez, Vincent K. Tsui, and S. J. B. Yoo, "A simplesimple, polarization independent, and dispersion insensitive SCM signal extraction technique for optical switching system using a fiber Bragg grating filter without fiber dispersion compensationapplications," submitted to *IEE Electronic Letters* (2001).

23. Jian Wang, Biswanath Mukherjee, S. J. Ben Yoo, 'MPLambdaS simulator', submitted to *IEEE Communications Magazine*

24. Wushao Wen, Biswanath Mukherjee, S. J. Ben Yoo, 'Sub-Path protection in MPLambdaS networks', to be submitted to *IEEE Journal of Selected Areas in Communications*

25. S. J. B. Yoo, Hyuek Jae Lee, Srikanth Vaidianathan, Katsunari Okamoto, Shin Kamei, "Optical-Label Switching and Routing by Tunable Wavelength Conversion and Uniform Loss Cyclic Frequency Array-Waveguide Grating," Technical Digest, *Optical Fiber Communication Conference*, Anaheim, California, paper WDD49 (2001).

26. H. J. Lee, S. J. B. Yoo, Changsoo Park, "Novel all-optical 10 Gbps RZ-to-NRZ conversion using SOA-loop mirror," Technical Digest, *Optical Fiber Communication Conference*, Anaheim, California, paper MB7 (2001).

27. Shun Yao, S. J. B. Yoo, and Biswanath Mukherjee, "A comparison study between slotted and unslotted all-optical packet-switched network with priority-based routing," Technical Digest, *Optical Fiber Communication Conference*, Anaheim, California, paper TuK2 (2001).

28. B. Meagher, G. K. Chang, G. Ellinas, Y. M. Lin, W. Xin, T. F. Chen, X. Yang, A. Chowdhury, J. Young, S. J. B. Yoo, C. Lee, M. Z. Iqbal, T. Rober, Y. J. Chen, and W. I. Way, "Design and Implementation of Ultra-low Latency Optical Label Switching for Packet-Switched WDM Networks," **18**, 1978, *IEEE/OSA Journal of Lightwave Technology* (2000).
29. Shun Yao, S. J. Ben Yoo, Biswanath Mukherjee, and Sudhir Dixit, 'All Optical Packet-switched Networks: A Study of Contention-Resolution Schemes in an Irregular Mesh Network with Variable-size Packets,' *to appear in SPIE journal Opticomm 2000.*
30. Shun Yao, S. J. Ben Yoo, Biswanath Mukherjee, and Sudhir Dixit, 'All Optical Packet-switched Networks and Data Networking,' *to appear in IEEE Communications Magazine* **(Invited)..**
31. S. J. Ben Yoo, 'Optical Label Switching and MPLS', *to be submitted to IEEE Networks*
32. Hyuek Jae Lee, S. J. B. Yoo, Vincent K. Tsui, and S. K. H. Fong, "A simple all-optical label detection and swapping technique incorporating for optical a fiber Bragg grating filter", *IEEE Photonic Technology letters* , vol 13, p. 635-637 (2001).
33. H. J. Lee, S. J. B. Yoo, Chang Soo Park, 'Novel all-optical 10 Gbp/s RZ-to-NRZ conversion using SOA-loop-mirror,' *to be submitted to IEEE Photon. Technol. Lett.*
34. Shun Yao, S. J. Ben Yoo, Biswanath Mukherjee, and Sudhir Dixit, 'A comparison study between slotted and unslotted all-optical packet-switched network with priority-based routing,' *submitted to IEEE JSAC.*
35. S. J. B. Yoo, Hyuek Jae Lee, Yanda Zhang, S. K. H. Fong, Vincent Hernandez, Vincent K. Tsui, Srikanth Vaidianathan, Katsunari Okamoto, Shin Kamei, "High-Speed All-Optical Packet Routing System with Optical-Label Switching and Optical-Label Swapping using Rapidly Tunable Wavelength Conversion and a Uniform-Loss Cyclic Frequency AWGR," submitted to *IEEE photonics technology letters* (2001).
36. S. J. B. Yoo, 'Wavelength Conversion and Crossconnect,' *to be submitted to Fiber and Integrated Optics (***Invited)**

Presentations-

1. Jing Cao, Zhong Pan, Minyong Jeon, Yash Bansal, Julie Taylor, Venkatesh Akella, Shin Kamei, Katsunari Okamoto, S. J. B. Yoo, "Cascaded Operation of an Optical Packet Routing System with Optical-Label Switching and 2R-regeneratioin," submitted to *European Conference on Optical Communication* , Kopenhagen, Denmark (2002).

2. S. J. B. Yoo, "Rapidly Switching Optical Packet Routing System with Wavelength, Time, and Space Domain Contention Resolution**,"** Technical Digest, *Photonics in Switching Conference 2002,* Cheju, Korea (2002) **(Invited).**

3. Jing Cao, Zhong Pan, Minyong Jeon, Yash Bansal, Julie Taylor, Venkatesh Akella, Shin Kamei, Katsunari Okamoto, S. J. B. Yoo, "Cascaded Operation of an Optical Packet Routing System with Optical-Label Switching and 2R-regeneration," accepted for publication in *European Conference on Optical Communication* , Kopenhagen, Denmark (2002).

4. S. J. B. Yoo, "All-Optical Switching: All-Optical Label Switching Approach", Optical Networking Workshop Panel Discussions, 7th OptoElectronics and Communications Conference, Yokohama, Japan (2002) **(Invited).**

5. S. J. B. Yoo, "All Optical Label Switching: Architecture, Protocol, and System Demonstration", to be published in Technical Digest, 7th OptoElectronics and Communications Conference, Yokohama, Japan (2002) **(Invited).**

6. S. J. B. Yoo, "Expermental Research Networks", CENIC 2002 **(Invited).**

7. Fei XUE, S. J. Ben Yoo, "Self-similar Traffic Shaping at the Edge Router in Optical Packet-switched Networks," *The IEEE International Conference on Communications 2002* (2002).

8. S. J. B. Yoo, Yash Bansal, Zhong Pan, Jing Cao, Vincent K. Tsui, Steven K. H. Fong, Yanda Zhang, Julie Taylor, Hyuek Jae Lee, Minyong Jeon, Venkatesh Akella, Katsunari Okamoto, Shin Kamei, "Optical-Label Switching based Packet Routing System with Contention Resolution Capabilities in Wavelength, Time, and Space Domains," *in Technical Digest, IEEE/OSA Optical Fiber Communication Conference*, paper #WO2, Anaheim, California, (2002).

9. Fei Xue, Shun Yao, Biswanath Mukherjee, and S. J. B. Yoo, "The performance improvement in optical packet-switched networks by traffic shaping of self-similar traffic," *in Technical Digest, IEEE/OSA Optical Fiber Communication Conference*, paper #WG4, Anaheim, California, (2002).

10. V.J. Hernandez, Z. Pan, J. Cao, V.K. Tsui, Y. Bansal, S.K.H. Fong, Y. Zhang, M.Y. Jeon, S.J.B. Yoo, Brian Bodtker, Steven Bond, William J. Lennon, Hank Higashi, Bryan Lyles, Robert McDonald, "First Field Trial of Optical Label Switching and Packet Dropping on a 477km NTON/Sprint Link," *in Technical Digest, IEEE/OSA Optical Fiber Communication Conference*, paper #TuY4, Anaheim, California, (2002).

11. Zhong Pan, Jing Cao, Yash Bansal, Vincent K. Tsui, Steven K. H. Fong, Yanda Zhang, Julie Taylor, H. J. Lee, Minyong Jeon, Venkatesh Akella, S. J. B. Yoo, Katsunari Okamoto, Shin Kamei, "All-optical programmable time-slot-interchanger using optical-label switching with tunable wavelength conversion and *N* by *N* arrayed waveguide grating routers," *in Technical Digest, IEEE/OSA Optical Fiber Communication Conference*, paper #WM6, Anaheim, California, (2002).

12. S. J. B. Yoo, "Demonstration of optical packet switching using optical-label switching technologies," SPIE Asia-Pacific Optical and Wireless Communications Paper 4582-21 (2001) (**invited).**

13. S. J. B. Yoo, "Optical-label switching, MPLS, MPLambdaS, and GMPLS," SPIE Asia-Pacific Optical and Wireless Communications Paper 4585-02 (2001) (**invited).**

14. S. Yao, S. J. B. Yoo, B. Mukherjee, S. Dixit, "Hybrid contention resolution for an optical packet-switched network with self-similar IP traffic," SPIE Asia-Pacific Optical and Wireless Communications Paper 4585-04 (2001).

15. W. Wen, B. Mukherjee, S. J. B. Yoo, "QoS based protection in MPLS controlled WDM mesh networks**,"** SPIE Asia-Pacific Optical and Wireless Communications Paper 4585-17 (2001).S. J. B. Yoo, Hyuek Jae Lee, Srikanth Vaidianathan, Katsunari Okamoto, Shin Kamei, "Demonstration of High-Speed Optical-Label Switching and Routing by Tunable Wavelength Conversion and Uniform Loss Cyclic Frequency AWGR," Technical Digest, *Optoelectronics and Communications Conference*, Sydney, Australia, paper 1-16, pp. 374-5 (2001)

16. S. J. B. Yoo, Hyuek Jae Lee, Srikanth Vaidianathan, Katsunari Okamoto, Shin Kamei, "High-Speed Optical-Label Switching and Routing by Tunable Wavelength Conversion and Uniform Loss Cyclic Frequency Array-Waveguide Grating," Technical Digest, *Photonic Switching Conference*, Monterey, California (2001)

17. H. J. Lee, S. J. B. Yoo, Changsoo Park, "Novel all-optical 10 Gbps RZ-to-NRZ conversion using SOA-loop mirror," Technical Digest, *Optical Fiber Communication Conference*, Anaheim, California, paper MB7 (2001).

18. Shun Yao, S. J. B. Yoo, and Biswanath Mukherjee, "A comparison study between slotted and unslotted all-optical packet-switched network with priority-based routing," Technical Digest, *Optical Fiber Communication Conference*, Anaheim, California, paper TuK2 (2001).

19. S. J. B. Yoo, Hyuek Jae Lee, Srikanth Vaidianathan, Katsunari Okamoto, Shin Kamei, "Optical-Label Switching and Routing by Tunable Wavelength Conversion and Uniform Loss Cyclic Frequency Array-Waveguide Grating," Technical Digest, *Optical Fiber Communication Conference*, Anaheim, California, paper WDD49 (2001).

7. Inventions and Patents

1. Fei Xue, Julie Taylor, Sung-Joo Ben Yoo, invention disclosure, "Edge Routers for Optical Label Switched Network," UC Doc 2002-232 (2002)

2. Sung-Joo Yoo, "Method and Apparatus For Optical Signal Processing Using Subcarrier Multiplexed Headers," U.S. Patent Application, Doc UC 2001-189 (2001)

8. Conclusions

This final report summarizes the entire progress made during the 27 month contract period for the project titled Ultra-Low Latency Multi-Protocol Optical Routers for the Next Generation Internet. The progress made during this contract period fulfilled all the required deliverables and significantly advanced the optical router technology. In particular, this project demonstrated the first all-optical label switching router, the first all-optical label swapping with negative penalty, the first edge-router for optical-label switching, the first cascaded operation of optical-label switching routers, and the first field trial of optical-label switching routers. The optical router has achieved the following key performance milestones.

- Sub-nanosecond (~600 psec) switching speed of the optical router based on rapidly tunable wavelength converter,
- Innovative contention resolution with wavelength-time-space domain switching aided by electronic buffering to achieve packet loss rate below 0.5 % at load 0.6.
- FPGA based arbitration and contention resolution with ultralow 250 nsec latency,
- Combination of the edge router and the optical label switching router demonstrating the first cascaded all-optical label switching IP over WDM with packet loss rate below 0.1 %.