Multi-terabit hybrid photonic switching and routing

Lakshman S. Tamil and Andrea Fumagalli
Dept. of Electrical Engineering, University of Texas at Dallas, Richardson, Texas 75080

ABSTRACT

Constructing multi-Terabit switches and routers entirely out of electronic cross-bars require multiple racks of switches and huge number of interconnects. The cost and power consumption of such a system is very high and has a limited scalability. The all-optical switching has the problem of buffering and processing and the construction of multi-terabit switches and routers based on all-optical switching is not feasible in the near future. However, based on a hybrid approach, using optical cross-bars with nanosecond switching time and large bandwidth and DWDM interconnects in conjunction with electrical memory, one can construct huge and highly scalable switches and routers.

Keywords: optical switching, terabit switching, terabit router, terabit cross-connect

1. INTRODUCTION

1.1 Network layers

The convergence of data, voice and video over IP and the increasing deployment of broadband and 3G wireless are increasing the backbone traffic tremendously. The backbone infrastructure of the future should be capable of transporting and switching/routing multi-terabits of data with carrier class reliability. The all-optical network was proposed as the ultimate solution to the growing network capacity. In a true all-optical network, all the transport as well as switching and routing are done in the optical domain. In the ultimate all-optical network envisioned, the service offering is also done in the optical domain. The advantages cited for the all-optical network is the reduction in the cost due to elimination of O-E-O conversion and the transparency of the optical domain to various data formats and data rates. The all-optical network is not a panacea as it is advocated to be. Although, WDM has proven to be a great boon to transport offering high scalability with cost effectiveness, the switching/routing of huge quantities of data still remains a challenge. All-optical packet switching/routing is not feasible as a practical means because of the great challenges it faces in areas such as buffering and header reading and the cost effectiveness of it is also questionable.

<table>
<thead>
<tr>
<th>IP/MPLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>GbE/SONET/SDH/GE</td>
</tr>
<tr>
<td>OCh (EXC/EADM)</td>
</tr>
<tr>
<td>OMS (OXC/OADM)</td>
</tr>
<tr>
<td>OTS (FXC)</td>
</tr>
</tbody>
</table>

Fig. 1. Network Layering

The success of Gigabit Ethernet technology for MAN and WAN applications and the emergence of Multi-protocol labelswitching (MPLS) as a successful means of traffic engineering is changing the network layering into one shown in Fig. 1(IP/GbE/SONET/WDM) as opposed to the classical IP/ATM/SONET/WDM. The IP/WDM, which is the ultimate
transport of IP directly over wavelengths has many unanswered questions that include the WDM adaptation layer, optical 3R, error correction, grooming, protection and performance monitoring. Also, a successful network layering should match well with the service offering. The emergence of Gigabit Ethernet as a ubiquitous service platform from access to MAN to WAN and the reasons stated above will pre-empt the emergence of IP/WDM at least for a while.

1.2. Switching granularity

In a network that processes large amount of data, one easy way of handling switching is to develop hierarchy of switching with increasing switching granularity (Fig.2). The lowest of granularity of switching in a network is an IP packet (ATM cells are not considered here), whereas the largest granularity of switching is a fiber. The intermediate switching granularities are band of wavelengths, wavelength, and sub-wavelength which is larger than an IP packet. The rate of dynamic switching of these granularities varies inversely as the size of the granularity so that an IP router should dynamically reconfigure its input to output mapping say hundreds of thousand of times every second whereas a fiber cross-connect should change its input to output mapping once in every six months or a year. The more dynamic is the switch, the more complex is the control plane. Looking at the intermediate switching granularity namely, band of wavelengths, wavelength and sub-wavelength, the requirement on band of wavelengths switching is more static than dynamic, the requirement on wavelength switching is also more static than dynamic until wavelength is a commonly
offered service. The dynamic wavelength switching in a network involves complex wavelength assignment schemes; the assignment of the same wavelength in all links for setting up a light path is an NP complete problem whereas the assignment of different wavelengths in different links for establishing a light path requires wavelength conversion.

The possible schemes for sub-wavelength switching can be broadly classified as electrical and optical schemes. The optical burst switching, optical CDMA switching, and optical sub-carrier switching fall under the optical scheme whereas STS-n cross-connect and EADM come under the electrical scheme. The optical CDMA switching and optical sub-carrier switching have problems in implementation. Lot of research has been conducted in these areas and the ultimate conclusion is that they are not feasible and cost effective for deployment. The optical burst switching is nothing but time division switching of a wavelength. The implementation of this involves complex control plane involving among other things complex signaling schemes, resource management and burst assembly. The complexity of the optical burst switching outweighs its usefulness.

The conclusion based on the discussions made in the previous paragraphs and the consideration of emerging service offering and link layer technology, feasibility and cost effectiveness of various switching technologies is that the next generation network at least in the near term will look more like the one shown in Fig.1 with electrical IP switching, electrical STS-n XC/ADM and limited static wavelength switching, all interconnected by WDM transport. The wavelength switching can be accomplished by ROADM and OXC. These network elements have been developed and can be deployed without much difficulty. OXCs as big as 1024x1024 based on MEMS have been demonstrated.

In this paper we describe a multi-terabit hybrid photonic switching and routing sub-system made up of a semiconductor optical amplifier (SOA) switch matrix capable of sub-wavelength switching. The SOA switch matrix has cross-points with bandwidth as high as 400 Gb/s. This high bandwidth of SOA cross-points in conjunction with its nanosecond switching capability can be exploited to yield sub-wavelength switching when wavelength-space-time switching is employed. A highly scalable multi-terabit switch with lower cost and lower power becomes feasible using this approach.

2. SWITCH ARCHITECTURE

Sub-wavelength switching and routing are done in the electronics domain and involve O-E-O conversion. The electronic switch fabrics have cross-points with 2.5 Gb/s bandwidth or in some cases 10 Gb/s bandwidth. Developing a multi-terabit cross-connect, switch or router using electrical cross-points alone involves multi-stage switching. The huge number of cross-points and the multi-stages of switching required in conjunction with O-E-O conversion at every stage, and the fiber interconnect between the various switch stages, make this approach more costly, more power consuming and less scalable.

The architecture of the switching system described here is composed of two building blocks, the Photonic Switch Core (PSC) and the Photonic Switch Edge (PSE) as shown in Fig. 3. In the 5.2 Tb/s switching system that we demonstrated there are 16 I/O shelves each with 320 Gb/s capacity. The PSC is a rearrangeably non-blocking space switch using semiconductor optical amplifiers (SOA) as cross-points. In the case demonstrated here, the PSC is a 16x16 switch matrix (see Fig. 4). The PSEs are connected to the PSC through DWDM fiber interconnect.

2.1. Photonic Switch Edge (PSE)

The PSE receives electrical signals as STS-48 SONET streams from the I/O modules and creates containers to be transmitted as an optical stream at 12.5 Gb/s towards the PSC; it also receives containers as an optical stream from the PSC and regenerates continuous STS-48 SONET electrical streams towards I/O modules. The PSE performs three functions. The first function is ‘grooming and wavelength mapping,’ in which STS-1 grooming (in the case of SONET)
or packet grooming (in the case of packet over SONET) and cross-connecting to the desired STS-48 output stream, which in turn is mapped onto a desired wavelength, is accomplished.

The second function is ‘container processing and timeslot mapping,’ in which de-byte interleaving of the output SONET stream from the electrical grooming switches and creation of containers to be transmitted towards the PSC is performed. The container processing function also includes conversion of the continuous stream received into burstmode stream. In our case, the container processing generates, for every 125µs frame, 192 containers of payload information mapped from the 192 STS-1s coming out of the grooming stage and 3 containers of performance monitoring information.

The container consists of payload along with header information describing the wavelength and timeslot and forward error correction codes (FEC). The format of the container is shown in Fig 5. Each container is mapped onto a timeslot in
an outgoing wavelength. These functions are achieved using a proprietary ASIC. The output of this stage is a two-dimensional mapping of data onto wavelengths and timeslots as shown in Fig. 6.

![Container format](image1)

**Fig. 5** Container format

![Wavelength and time slot mapping](image2)

**Fig. 6** Wavelength and time slot mapping of the input to the SOA switch matrix

The third function of the PSE is the multiplexing of the containerized output into a high-speed electrical stream and transmitting over an optical wavelength. The data rate per wavelength is 12.5 Gb/s, which is 10Gb/s of payload with overheads including FEC. All wavelengths are multiplexed into one fiber and connected to PSC using a proprietary DWDM interconnect. In the case discussed here, there are 36 wavelengths per PSE; the extra four wavelengths take care of the additional bandwidth required for non-blocking interconnection.

### 2.2. Photonic Switch Core (PSC)

The Photonic Switch Core (PSC) is interconnected with all the PSEs (16 of them in the case discussed here). The PSC receives composite DWDM burst signals (see Fig. 7) from all the PSEs. The PSE is configured (by the two-dimensional wavelength-time mapping of the data described previously) in such a way as to switch each composite signal (multiple wavelengths in the same time slot) from one PSE to any other PSE during every time slot as per the memory mapping.
dictated by the cross-connect or router function. The fast SOA switching time (as low as few nanoseconds) enables fast reconfiguration of the PSC switch matrix. In the case discussed here, the matrix is reconfigured 195 times in a 125 μs period. The gap time between the reconfiguration (72 ns) is more than sufficient to overcome transient effects and includes training pattern necessary for burst mode receivers. The SOA switch matrix developed is shown in Fig. 8 below.

![Fig. 7. Time-space switching of SOA matrix](image)

![Fig. 8. Semiconductor Optical Amplifier (SOA) switch matrix](image)

A PSC made up of 16x16 SOA switch matrix, using 36 wavelengths per 320 Gb/s of PSE can switch non-blockingly any STS-1 to any other STS-1 among 98,304 STS-1s, providing a total switching capacity of 5.12 Tb/s. This can be scaled gracefully up to 40 Tb/s and beyond.

3. CONCLUSIONS

A highly scalable sub-wavelength switch of multi-terabit capacity can be built using semiconductor optical amplifiers (SOA) as cross-points. The sub-wavelength switching is achieved using the high-bandwidth and short switching time
properties of SOA. The strong advantage of optics over electronics lies in its DWDM property and not in its switching or processing ability. Here we have capitalized on the strength of optics and disguised its application for switching rather than for transport. The advantages of this approach include low power and low cost in addition to graceful scalability.

4. REFERENCES


Acknowledgements
This work was carried out when Lakshman Tamil was with Yotta Networks, Inc. Helpful discussions and support of Yotta Engineering staff and financial support of investors are gratefully acknowledged.