Introducing optical switching in high-capacity commercial routers

Belen Garcia-Manrubia, Pablo Pavon-Marino
Technical University of Cartagena, (UPCT), Plaza del Hospital 1, 30202, Cartagena, Spain
Tel: (34) 968 325952, Fax: (34) 968 325973, e-mail: pablo.pavon@upct.es

ABSTRACT
High-performance packet switches are costly, bulky and power-hungry. Their scalability is constrained by the limits of electronic technology at high binary rates. In this paper we study the benefits of introducing optical technology in the switch fabric architecture, to cope with this bottleneck. In particular, we study the replacement of the switch fabric boards in the Cisco 12816 architecture, by two optical architectures based on arrayed waveguide gratings (AWGs) and tunable wavelength converters (TWCs). The results illustrate the potential benefits in terms of power consumption and system scalability that these optical architectures could bring to commercial routers.
Keywords: optical packet switching, optical switching fabrics, wavelength-routing optical fabric.

1. INTRODUCTION

Electronic routers in the backbone of the Internet process millions of packets per second (Mpps). However, these high-performance packet switches present critical limitations. For instance, each generation of commercial routers is more power-hungry than the previous one, and it is more difficult to package them in a single chassis. As an example we can see the high-capacity router proposed by Cisco, CRS1, which consists of 80 chassis and weighs up to 700 kg per chassis. To overcome the electronic bottleneck, a significant number of solutions have been proposed addressing the power consumption and scaling problems introducing optical technology in the routers. In this paper the alternative studied is implementing some specific router functionalities using optical technology. In particular, following the conclusions achieved in [1] we decide to implement an optical switching fabric, while keeping the packet processing and buffering in its electronic form. Therefore, we focus on hybrid high-performance packet switches with optical fabrics. We hypothesize that this approach can potentially scale better to higher capacities, and at the same time significantly reduce the power consumption.

During the last decade many packet switch architectures using optical switching fabrics have been studied such as [2]-[4]. However, they are often far from being commercially feasible, since they require optical components which are still either in their beginnings or, simply, too expensive. As a result, during the last years researchers have changed their strategy paying more attention to optical architectures which require optical components that are commercially available [5]. Different proposals of optical switching fabric architectures have been developed using diverse optical elements such as array wavelength gratings (AWGs) [5], [6].

In this paper we opt to use an AWG-based architecture because of several reasons. Firstly, this architecture is designed using off-the-shelf components, differently from many works appeared previously in the literature. Moreover, AWG is a promising passive device because it is able to scale to ultra-high capacities while ensuring significant power savings [7]. Our objective is to show that if switch fabrics used in commercial routers are replaced by optical switch fabrics, important advantages in terms of scalability and low power consumption can be met. To this end we choose a VOQ (Virtual Output Queuing) router (Cisco 12816) whose electronic switch fabric architecture is clearly identifiable, suggesting that a substitution by an optical switch fabric could be possible. Then, we show the results obtained for this router as an example of the potential benefits achievable.

2. RELATED WORK: OPTICAL SWITCH FABRICS

During the last decades, two main approaches of optical interconnection architectures have been considered: broadcast-and-select and wavelength routing switch fabrics [2]-[6]. A basic broadcast-and-select switch fabric consists of basically two stages. In the first stage, packets from all input ports, each on a different wavelength, are combined within the switch and are broadcast to all the output ports. In the second stage, wavelength selectors are used at each output port to select a wavelength, and consequently a packet, to be switched. On the other hand, a wavelength routing switch fabric (Figure 1) is based on an array wavelength grating (AWG) which is used to route the packets to the correct output ports by tuning the transmitters to the proper wavelength associated to the fixed receiver.

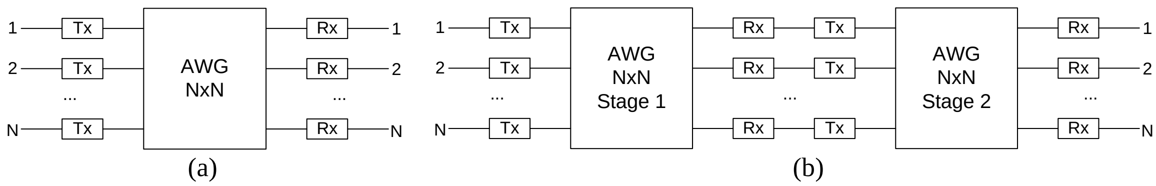

Figure 1. Wavelength routing architectures: AWG-based (a) one-stage and (b) two-stage switch architectures.

Broadcast-and-select architectures present advantages as a natural implementation of multicasting, but also disadvantages as the inevitable losses due to the broadcast stage. In its turn, wavelength routing architectures offer a good behavior respect to line rate variations and low power consumption. However, their main limitation is the in-band crosstalk introduced by the AWG. This effect appears when several input ports are using the same wavelength.

While both types of architectures present pros and cons, recently the wavelength routing approach seems to be overtaking the broadcast-and-select strategy, since research focused on AWG-based architectures have made interesting progress. In Figure 1(b) we present the architecture proposed in [8] to overcome the in-band crosstalk limitation by introducing a second stage. According to [8] it is proved that this two-stage switch can avoid the in-band crosstalk limiting the maximum number of ports using the same wavelength to four. Therefore, this limit will allow us to build large-size switching fabric without major crosstalk constraints.

3. CASE STUDY: CISCO 12816

In this section our objective is to evaluate if the switch fabrics used in commercial router Cisco 12813 [9] can be replaced by an optical switch fabric, achieving advantages in terms of space, throughput scalability, and power consumption. Cisco 12816 router [9] was chosen since its switch fabrics are memory-less, and thus if they are replaced by a memory-less optical fabric, scheduling processes in the switch would be unaffected.

3.1 Cisco 12816 description

Cisco 12000 Series is a portfolio of routers which deliver capacity and services with a distributed forwarding architecture and a crossbar switch fabric. Specifically, we study Cisco 12816 since this is the router with the highest aggregate switching capacity (1280 Gbps) in this series. Cisco 12816 router consists of an only chassis which contains 16 slots with a full-duplex throughput of 40 Gbps/slot. Some general data of this chassis are collected in Table I.

Table I. Cisco 12816 specifications.

Parameter	Cisco 12816
maximum weight (kg)	177
physical dimensions (WxHxD) (cm)	43.8 x 181.6 x 55.9
throughput (Gbps)	1280
power consumption (W)	4800

Cisco 12816 router (Figure 2) is divided into two main parts:

- *Backplane*: Gigabit Route Processors (GRPs) and Line Cards (LCs) are installed from the front of the chassis and plug into a passive backplane. This backplane contains serial lines that interconnect all of the line cards to the switch fabric cards, as well as other connections for power and maintenance functions.
- *40 Gbps Switch Fabric* (1280 Gbps switching system bandwidth): the core of the router is a crossbar switch fabric which consists of 2 clock scheduler cards (CSCs) and 3 switch fabric cards (SFCs). One CSC and the three SFCs are the active switch fabric; the second CSC provides redundancy for the other 4 cards. The SFC receives the scheduling information and clocking reference from the CSC, and performs the switching functions.

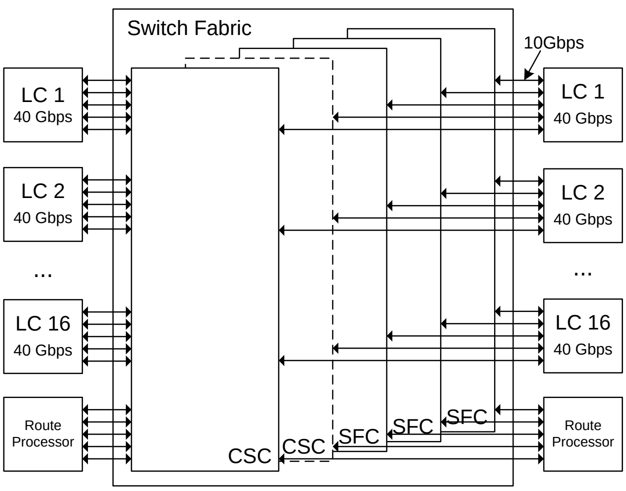

Figure 2. Cisco 12000 Series Internet Router.

The switch fabric can be imagined as an *N*x*N* non-blocking crossbar switch fabric where *N* stands the maximum number of LCs that can be supported in the chassis including the GRP (*N*=16). Each LC has *N*+1 virtual output queues (VOQ); one for each possible line card destination and one for multicast. The elements which form the switch fabric (see Fig. 2) are 2 clock scheduler cards (CSCs) and 3 switch fabric cards (SFCs). In our study, we assume that each of these five cards is replaced by an optical switch fabric. Naturally, the system synchronization functionalities CSC cards provide, that the replacing optical switch fabrics do not implement, are assumed to be provided by other hardware in the system.

3.2 Cisco 12816: Switch fabric replacement results

In this subsection, consumption and throughput data are studied when each of the five switch fabric cards in the Cisco 12816 (Figure 2) are substituted by the two AWG-based switching architectures considered (Figure 1). In summary, we investigate two cases: (i) an optical switch fabric is a one-stage AWG architecture, where the worse case in-band crosstalk occurs if all the TWCs are tuned in the same wavelength, (ii) a two-stage AWG-based optical switch fabric is used, applying the technique presented in [8], so that the worse case in-band crosstalk occurs if four TWCs are tuned in the same wavelength.

In order to know when these AWG-based architectures are feasible according to the physical impairments we use the data included in [5]. On the other hand, to obtain the power consumption data we use the following values: the power supply for a fixed laser is 2.5mW/Gb/s [10], the total power consumed by the fixed receiver is 1.1mW/Gb/s [11], and the electronic power required for tuning operation requires approximately 10mW [12].

In Table II consumption data and component count of the two substitution cases are collected. This table is divided into two parts. First, we start focusing on the so-called 16x16 solution. For this case, we try to build 16x16 optical fabrics that exactly match the throughput and size of the original Cisco architecture. The power consumption values obtained for the optical systems are 5.6 W and 11.3 W, in the one and two stages configuration respectively (both of them feasible from the signal degradation point of view). By comparing these values with the consumption of the original switch fabrics (807 W, 2 CSCs x 177 W/card, and 3 SFCs x 151 W/card) we see a 98%-99% of power consumption reduction.

Table II. AWG-based Architecture Consumption and Throughput Data for Cisco 12816

	16x16 Solution			N_{max}xN_{max} Solution	
Optical Architecture	One-stage	Two-stage		One-stage	Two-stage
no. NxN AWGs	5 16x16	10 16x16		5 16x16	10 127x127
no. Txs and Rxs	160	320		160	2540
throughput (Gbps)	1280	1280		1280	10160
total no. slots	16	16		16	127
consumption (mW)	5680	11360		5680	73520

The second part of Table II (N_{max}xN_{max} Solution) studies the case in which we try to build optical fabrics as large as possible in number of 40 Gbps ports, while still feasible from the signal impairments point of view. In other words, we try to assess the benefits in throughput increase that could be obtained by replacing the electronic switch fabrics by their optical counter-parts.

If we focus on the one-stage solution, we can see data are equal to data for one-stage in the first part of Table II. This is due to the maximum port account coincides with the original number of LCs (16). When we consider the two-stage solution, the number of line cards increases dramatically from 16 to 127 leading a throughput growth from 1280 to 10160 Gbps. Interestingly, the accumulated consumption values of the five 127x127 optical fabrics are still reduced (73 W), even if we refer to the original data (807 W). Therefore, two-stage AWG-based optical switch fabrics provide a great improvement in switch fabric throughput scalability, while still keeping significantly low values of power consumption.

4. CONCLUSIONS

In this paper we study the benefits of introducing optical technology in the switch fabric architecture, to cope with the electronic bottleneck. In particular, we study the replacement of the switch fabric boards in the Cisco 12816 architecture, by two optical architectures (Figure 1) based on arrayed waveguide gratings (AWGs) and tunable wavelength converters (TWCs). The results obtained illustrate the potential benefits in terms of power consumption and system scalability that these optical architectures could bring to commercial routers. Specifically, consumption savings up to 99% can be achieved with the same system throughput, or eight time increases in the throughput can be obtained while still consuming a ~10% fraction of the power of the electronic fabrics. These results suggest that introducing optical technology in the switch fabric architecture of high-performance routers can become an effective solution to overcome the electronic bottleneck according to the results achieved.

ACKNOWLEDGEMENTS

The work described in this paper was carried out with the support of the BONE project ("Building the Future Optical Network in Europe"); a Network of Excellence funded by the European Commission through the 7th ICT-Framework Program. This research has been supported by the MICINN/FEDER project grant TEC2010-21405-C02-02/TCM (CALM), it is also developed in the framework of the projects from Fundación Seneca 00002/CS/08 (FORMA) and "Programa de Ayudas a Grupos de Excelencia de la Región de Murcia", F. Séneca (Plan Regional de Ciencia y Tecnología 2007/2010).

REFERENCES

[1] R. S. Tucker, "The role of optics and electronics in high-capacity routers," *IEEE Journal of Lightwave Technology*, vol. 24, pp. 4655–4673, Dec. 2006.

[2] H.J. Chao et al., "Terabit/s ATM switch with optical interconnection network," *IEEE Lasers and Electro-Optics Society*, vol. 1, pp. 130–131, 1997.

[3] Y. Maeno, Y. Suemura, A. Tajima, and N. Henmi, "A 2.56-Tb/s multiwavelength and scalable fabric for fast packet-switching networks," *IEEE Photonic Technology Letters*, vol. 10, no. 8, pp. 1180–1182, 1998.

[4] H. Jonathan Chao and Ti-Shiang Wang, "An Optical Interconnection Network for Terabit IP Routers," *IEEE Journal of Lightwave Technology*, vol. 18, no. 12, pp. 2095–2112, 2000.

[5] R. Gaudino, G. A. Gavilanes, F. Neri, and J. M. Finochietto, "Can Simple Optical Switching Fabrics Scale to Terabit per Second Switch Capacities?," *IEEE Journal of Optical Communications and Networking*, vol. 1, no. 3, pp. B56-B69, Aug. 2009.

[6] F. Neri et al., "Capacity/Cost Tradeoffs in Optical Switching Fabrics for Terabit Packet Switches", ONDM 2009, Braunschweig, Feb. 2009.

[7] E. Bonetto, D. Cuda, G. A. Castillo, and F. Neri, "The role of arrayed waveguide gratings in energy-efficient optical switching architectures," OFC/NFOEC'10, San Diego, CA, Mar. 2010.

[8] A. Bianco, D. Hay, and F. Neri, "Crosstalk-Preventing Scheduling in Single- and Two-Stage AWG-Based Cell Switches," *IEEE Transactions on Networking*, vol. 19, issue 1, pp. 142–155, 2011.

[9] http://www.cisco.com/en/US/products/hw/routers/ps167/ps5632/index.html, (last access: 12th April, 2011)

[10] C. Kromer, G. Sialm, C. Berger, T. Morf, M. L. Schmatz, F. Ellinger, D. Erni, G.-L. Bona, and H. Jackel, "A 100-mW 4 x 10 Gb/s transceiver in 80-nm CMOS for high-density optical interconnects," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2667–2679, 2005.

[11] S. J. Koester and C. L. Schow and L. Schares, G. Dehlinger, J. D. Schaub, F. E. Doany, and R. A. John, "Ge-on-SOI-Detector/Si-CMOS-Amplifier Receivers for High-Performance Optical-Communication Applications," *IEEE Journal of Lightwave Technology*, vol. 25, no. 1, pp. 46–57, 2007.

[12] D. Hohlfeld, M. Epmeier and H. Zappe, "A thermally tunable, silicon based optical filter," *Sensors and Actuators A: Physical*, vol. 103, no. 1-2, pp. 93–99, 2003.