

Applying Optical Reconfiguration on ATM Switch Fabrics

H. S. Laskaridis, G. I. Papadimitriou, Senior Member, IEEE, and A. S. Pomportsis, Member, IEEE
Dept. of Informatics, Aristotle University of Thessaloniki
P.O. Box 888, 54124, Thessaloniki, Greece

Abstract

Altering the internal structure of an ATM switch fabric, based on the correlation among ports, can be proved to be advantageous in terms of performance, especially in LAN or campus ATM switches, where we witness stronger traffic correlation. Such reconfiguration can be easily performed in the optical domain, using simple optical elements. We prove the performance improvement, by applying data collected from a campus production ATM switch onto our proposed architecture.

Index Terms — ATM, switch fabric, reconfiguration, correlation

1. Introduction

The exponential increase in Internet's users and usage urged network engineers and researchers to search for ways to take advantage of the extreme transmission capacity of optical fiber as a physical medium and bridge the huge gap between the capabilities of fiber as a medium and the capabilities of terminating, switching and routing equipment. In [1] terms "optical cross-connections", "optical switching" and "optical routing" are clarified. Devices that are currently available in production networks, i.e. programmable optical cross-connects (OXC) and optical add/drop multiplexors (OADM), are able to multiplex and switch light wavelengths, but they are unable to do switching on a per packet basis.

There is a number of serious problems in optics, described in [1-3], concerning buffering (lack of optical RAM – delay lines are used instead), packet synchronization, optical crosstalk between paths sharing the same switching elements in optical multistage interconnection networks, path-dependent loss, and network management (basically troubleshooting). Most of them are not applicable to optical cross-connects or even switching, but are the reason for making optical routing currently impossible.

In a number of papers regarding **photonic packet switches** (e.g. [4-8], [18]), proposing per packet switching, the following approach is described: the header of each packet is located, extracted and converted to electrical

signal in order to be processed. The switch fabric is properly electronically reconfigured, based on the header's information, in order for the payload to be switched in the optical domain, without O-E-O conversion. In [9-10] a number of open issues regarding photonic packet switches (e.g. contention resolution, header's information extraction without O-E conversion), along with possible future solutions, are presented. In [11] a photonic switch that performs header extraction, recognition and replacement, all in the optical domain, is presented. However in [12] it is noted that "since optical technology is still in its early stage, optical processing is not yet a mature and viable alternative to the electronic processing of packet header".

On the other hand there are standards (e.g. ITU Recommendation I.311 [13] and ITU Recommendation I.610 [14]) that have to be supported, which seems to be very difficult in the optical domain. Moreover existing QoS support algorithms are based on buffer management algorithms. These algorithms cannot be easily implemented on optical buffers, i.e. delay lines, making QoS support in the optical domain cumbersome.

Thus, given the present immaturity of optical technology, we consider O-E-O conversion as mandatory. An all-optical switch, executing all control and maintenance functions in the optical domain, is not foreseeable in the near future.

In this paper we use currently available optical technology to propose an enhancement of a switch fabric presented in our previous work [15], namely **Grid-based ATM Switch Architecture (GASA)**. The new enhanced version takes advantage of the strong correlation between ATM ports. A passive star coupler, a relatively simple optical element, is used in order for the switch to be easily reconfigured in the optical domain. Passive star couplers are included in a number of proposed optical switches' architectures (for instance [16-19]).

The rest of the paper is organized as follows: In section 2 we demonstrate the correlation between ATM ports. We use our university campus ATM network as an example. In section 3 we present the Reconfigurable GASA architecture. In section 4 the performance improvement gained by applying reconfiguration is demonstrated using actual data. Finally, section 5 concludes the paper.

2. The concept of correlation

As we are moving from core backbone networks to campus or local networks the concept of traffic correlation between ports on an ATM switch becomes stronger and more obvious. Apparently, models using traffic distribution tables where each input port has equal probability to transmit towards each output port is far from realistic. In reality, there are pairs of input and output ports that cells never move from one to another, merely because the end stations connected on these ports never communicate. On the other hand there may be an end station that transmits all its traffic or a large portion of it towards one and only one end station. So basically in this paper we consider “large-time-scale” correlation, in contrast to “short-time-scale” correlation, i.e. bursts of cells created by the segmentation of larger (e.g. IP) packets.

The graphs presented in figure 1, created using Multi-Router Traffic Grapher (MRTG), reveal the correlation between two pairs of ATM ports from our University’s campus network. Each graph shows the incoming (shaded) and outgoing (line) traffic rate (in bits per second) of an ATM port for approximately 33 hours.

Although the graphs do not contain information about source and destination of cells, the correlation between the two pairs of ports [(a)-(b) and (c)-(d)] is obvious: note that outgoing traffic presented in graphs (a) and (c) is almost identical to incoming traffic presented in corresponding graphs (b) and (d) and vice versa.

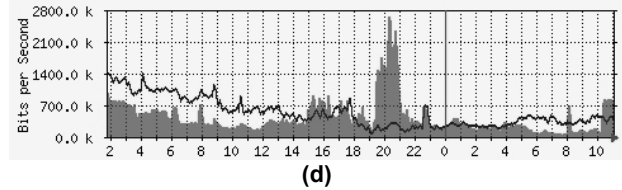
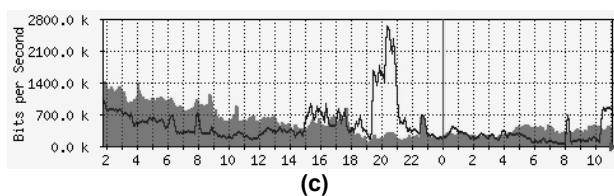
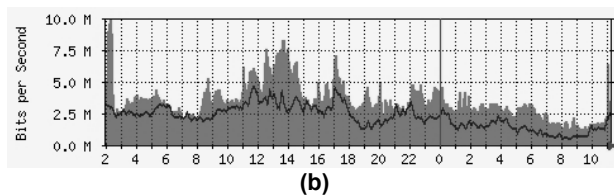
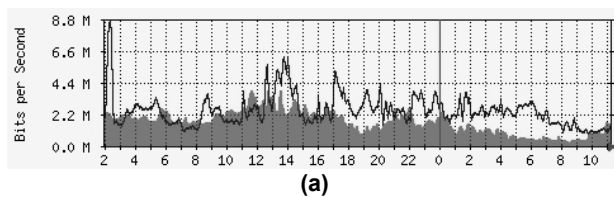


Figure 1: Daily traffic statistics of various ATM ports

The problem is how to take advantage of this correlation in the design of an ATM switch, given that the correlation depends on the end stations that the user connects on the switch, which port is used for each end station, and even the applications and services run on each end station.

3. Re-configurable GASA architecture

In [15] we have presented the Grid-based ATM Switch Architecture (GASA). In a GASA switch an input port and an output port are connected on a switching element (SE). The SE’s are interconnected forming a grid. Proper addressing of the SE’s makes the implementation of a simple cell routing algorithm feasible.

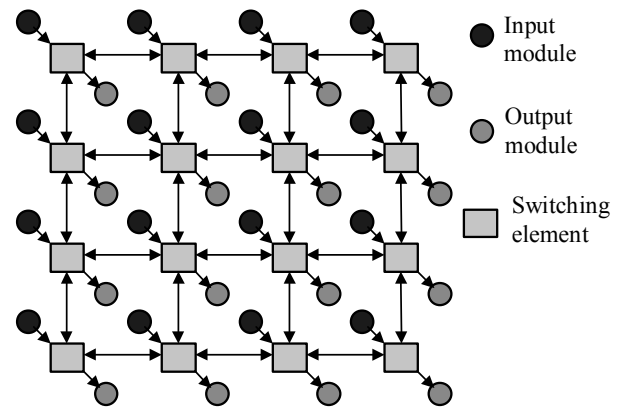


Figure 2: Grid-based ATM Switch Architecture (GASA)

Unlike other well-known architectures, such as Banyan and Clos networks, all cells in a GASA switch do not traverse paths of equal length. So it would be advantageous to connect inputs and outputs with high traffic among them on the same SE or on SE’s that are close in the grid.

In order to address the problem of taking advantage of the traffic correlation, we add an intermediate stage between input ports and SE’s, before O/E takes place. This stage consists of a tunable wavelength converter (TWC) for each input port, a passive star coupler and a fixed optical filter for each SE. Traffic statistics, regarding source and destination of each cell, are collected and transmitted to a control unit, namely Wavelength Assign-

ment Controller (WAC). Periodically WAC executes the “stable marriages” algorithm [20], using the Gale-Shapley method, in order to find the “best matching” between input and output ports. “Best matching” in this case is defined as the matching that minimizes the number of cells that have to go through more than one SE. According to the “stable marriages” algorithm, each input and output port has a sorted “preference list”, defining the “amount of satisfaction” that each port gets when connected to another port. The preference list is easily extracted from the actual traffic matrix, which is created by collected measurements. The “stable marriages” algorithm maximizes the overall “satisfaction”.

tion algorithm can be found in the following section.)

4. Performance improvement

The campus ATM network of Aristotle University of Thessaloniki (AUTH) consists of 4 ATM switches. The border switch connects AUTH network to the Internet and has 24 optical ports. Apart from the connection to our internet service provider and the interconnection to two other backbone switches, this switch has the following elements connected on it: 2 backbone routers, the main mail and DNS server of AUTH-net, the main web, proxy and FTP server of AUTH-net, a number of ATM worksta-

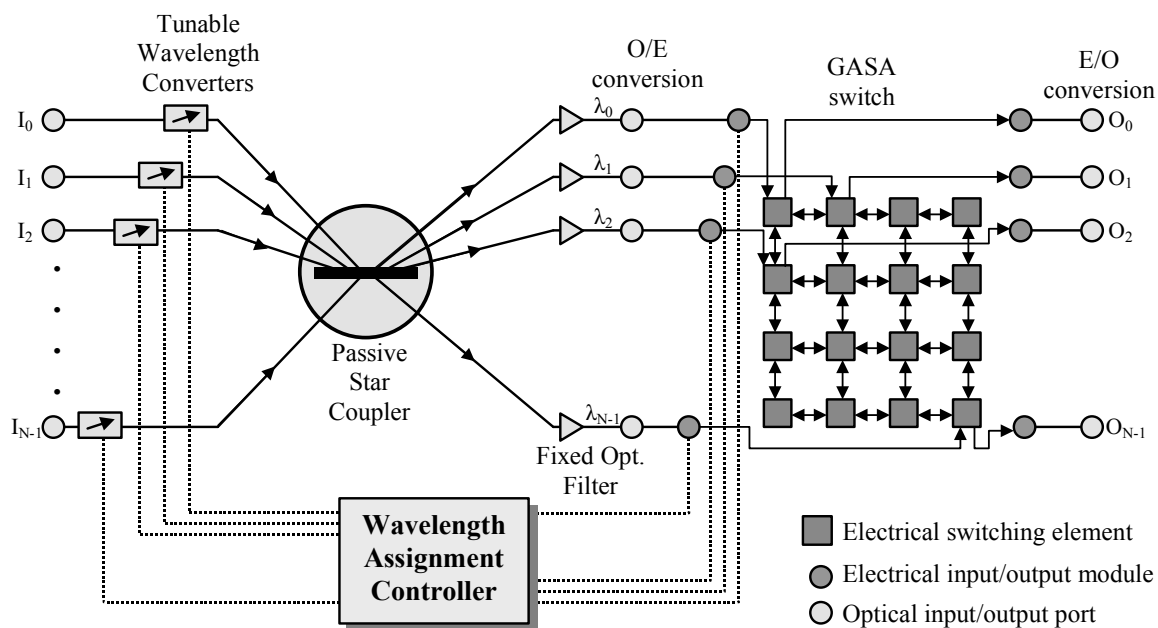


Figure 3: Re-configurable Grid-based ATM Switch Architecture

Thus, based on the traffic seen by the switch during a previous time period and assuming that the traffic characteristics of the new time period will be the same, the connections of input ports with the grid change. If the statistics show that it is advantageous to connect input port I_x with output port O_y (through SE_y of course), WAC orders the TWC connected to I_x to transmit in wavelength λ_y . In this way optical signal coming from I_x will be received by the optical filter connected with SE_y . O_y is also connected to SE_y , thus having high probability of switching the cell without crossing other SE's.

Thus, the “broadcast-and-select” architecture is used to make the switch re-configurable. We note that “stable marriages” algorithm has complexity $O(N^2)$, where N is the number of ports. However this is of minor importance, as reconfiguration does not take place very frequently. (Comments on the frequency of executing the reconfigura-

tions and fast ethernet switches with ATM up-links.

We have collected data about the traffic switched through the border switch for 2 regular working days. We calculated the distribution of cells switched during the second day (approximately 4.6×10^{12} cells) based on the number of SE's that they would have to cross, in order to reach their destination SE, if the switch was a GASA switch.

The distribution of cells was calculated for the following cases:

1. if no reconfiguration takes place,
2. if reconfiguration takes place, based on the statistics of
 - a) the previous day,
 - b) the previous 10-minutes period,
 - c) the corresponding 10-minutes period of the previous day.

Results are presented in figure 4. In case (1) we had average number of hops 3.530, in case (2a) 2.26, in case (2b) 2.12 and in case (2c) 2.29. However, in order to evaluate the methods of reconfiguration, the portion of cells transmitted with a single hop (i.e. crossing only one SE) should be used as a metric instead of the average number of hops. This is because the reconfiguration algorithm attempts to

low that cannot pause any performance restrictions, regarding the operation of WAC.

5. Discussion, conclusions and future work

Unfortunately, traffic traces with exact timestamps on a per-cell basis are not available due to technical reasons.

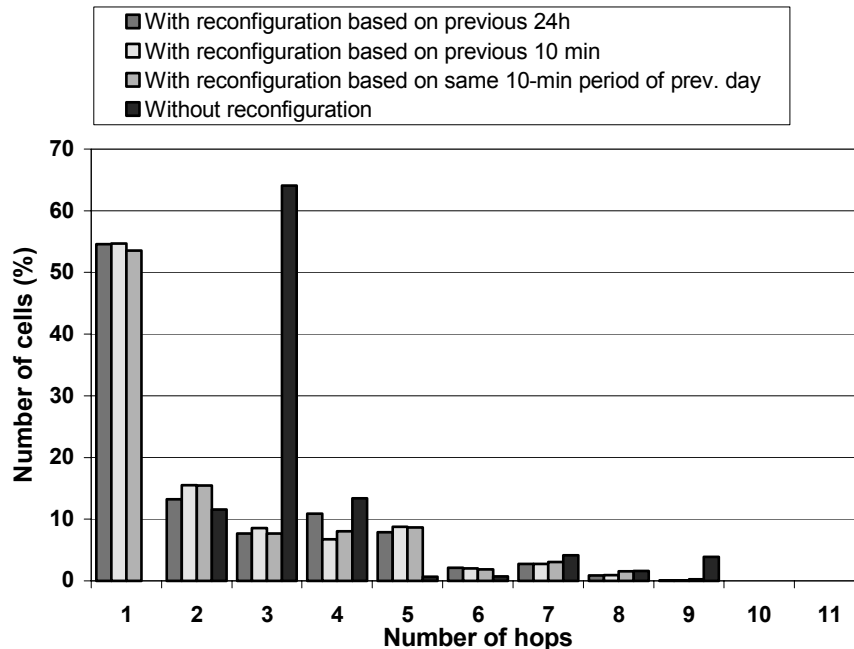


Figure 4: Distribution of cells based on number of hops, with or without reconfiguration

maximize the number of cells switched with a single hop, by connecting pairs of input and output ports with high traffic on the same SE. Reconfiguration based on the previous 10-minutes period performs better than other reconfiguration schemes, although differences are marginal: in our experiment, the portions of cells that were switched crossing only one SE were:

- 54.691% when reconfiguration based on previous 10 minutes was applied.
- 54.590% when reconfiguration based on previous 24 hours was applied.
- 53.580% when reconfiguration based on the corresponding 10-minutes period of the previous day was applied.

We note that cell distributions presented in figure 4 do not follow any of the standard distributions, because they depend on the need for communication between the end stations connected to each port. The actual 24 ports of the border switch were mapped on ports 0 to 23 of a simulated 32-port GASA switch (with grid diameter equal to 11), keeping the order intact.

Note also that the frequency of executing the “stable marriages” algorithm (maximum every 10 minutes) is so

Such traces would enable us to evaluate other performance metrics as well, such as delay and loss probability. However, we can intuitively expect that increasing the number of cells being switched without crossing more than one SE, results in decongesting the inter-SE’s links, thus decreasing cell loss. Average delay is also expected to decrease, as fewer cells are queued in multiple queues of multiple SE’s.

Making the “output side” of the switch fabric also reconfigurable would improve the performance even more. Arranging the connections of both input and output ports on the grid of SE’s in such a way that the average number of hops would be minimal, would result in further performance improvements. However this comes on the expense of a much more sophisticated algorithm.

The idea of applying optical reconfiguration prior to main switch fabric could also be applied to well-known architectures such as Banyan networks. A different wavelength assignment algorithm, taking under consideration the structure of the Banyan network, could decrease the blocking probability and consequently the amount of necessary internal buffers.

The idea of taking advantage of correlation, based on

statistics collected over a previous time period, could be deployed even in routing protocols. For instance a modified PNNI protocol could make decisions regarding the path selection while establishing new ABR or UBR virtual circuits, based on the amount of anticipated traffic.

References

- [1] L. Thylen, G. Karlsson, O. Nilsson, "Switching Technologies for Future Guided Wave Optical Networks: Potentials and Limitations of Photonics and Electronics", *IEEE Communications Magazine*, Vol.34, No.2, 1996, pp.106-113
- [2] Yi Pan, C. Qiao, Y. Yang, "Optical Multistage Interconnection Networks: New Challenges and Approaches", *IEEE Communications Magazine*, Vol.37, No.2, pp.50-56, 1999
- [3] B. Mukherjee, "WDM Optical Communication Networks: Progress and Challenges", *IEEE Journal on Selected Areas in Communications*, Vol.18, No.10, pp.1810-1824, 2000
- [4] S. Danielsen, B. Mikkelsen, C. Joergensen, T. Durhuus, K. Stubkjaer, "WDM Packet Switch Architectures and Analysis of the Influence of Tunable Wavelength Converters on the Performance", *IEEE Journal of Lightwave Tech.*, Vol.15, No.2, pp.219-226, 1997
- [5] S. Danielsen, C. Joergensen, B. Mikkelsen, K. Stubkjaer, "Analysis of a WDM Packet Switch with Improved Performance Under Bursty Traffic Conditions Due to Tuneable Wavelength Converters", *IEEE Journal of Lightwave Tech.*, Vol.16, No.5, pp.729-735, 1998
- [6] D. Hunter, D. Cornwell, T. Gilfedder, A. Franzen, I. Andonovic, "SLOB: A Switch with Large Optical Buffers for Packet Switching", *IEEE Journal of Lightwave Tech.*, Vol.16, No.10, pp.1725-1736, 1998
- [7] W. Zhong, R. Tucker, "Wavelength Routing-Based Photonic Packet Buffers and Their Applications in Photonic Packet Switching Systems", *IEEE Journal of Lightwave Tech.*, Vol.16, No.10, pp.1737-1745, 1998
- [8] J. Diao, P. Chu, "Analysis of Partially Shared Buffering for WDM Optical Packet Switching", *IEEE Journal of Lightwave Tech.*, Vol.17, No.12, pp.2461-2469, 1999
- [9] S. Yao, B. Mukherjee, S. Dixit, "Advances in Photonic Packet Switching: An Overview", *IEEE Communications Magazine*, Vol.38, No.2, pp.84-94, 2000
- [10] D. Hunter, I. Andonovic, "Approaches to Optical Internet Packet Switching", *IEEE Communications Magazine*, Vol.38, No.9, pp.116-122, 2000
- [11] K. Park, T. Mizumoto, "Packet Header Recognition Assigning the Position of a Signal in the Time Axis and its Application to All-Optical Self-Routing", *IEEE Journal of Lightwave Tech.*, Vol.19, No.8, pp.1076-1084, 2001
- [12] V. Elek, A. Fumagalli, G. Wedzinga, "Photonic Slot Routing: A Cost-Effective Approach to Designing All-Optical Access and Metro Networks", *IEEE Communications Magazine*, Vol.39, No.11, pp.164-172, 2001
- [13] ITU Recommendation I.311: "B-ISDN General Network Aspects", Helsinki, Aug. 1996
- [14] ITU Recommendation I.610: "B-ISDN Operation and Maintenance Principles and Functions", Geneva, Nov. 1995
- [15] H. Laskaridis, A. Veglis, G. Papadimitriou, A. Pomportsis, "Grid-based switch fabrics: a new approach in designing fault-tolerant ATM switches", *Elsevier Computer Communications*, Vol.24, Issues 15-16, pp.1589-1606, 2001
- [16] G. Papadimitriou, D. Maritsas, "WDM passive star networks: a learning automata-based architecture", *Elsevier Computer Communications*, Vol.19, pp.580-589, 1996
- [17] G. Papadimitriou, D. Maritsas, "WDM star networks: hybrid random access and reservation protocols with high throughput and low delay", *Elsevier Computer Networks and ISDN Systems*, Vol.28, pp.773-787, 1996
- [18] A. Misawa, M. Tsukada, "Broadcast-and-Select Photonic ATM Switch with Frequency Division Multiplexed Output Buffers", *IEEE Journal of Lightwave Tech.*, Vol.15, No.10, pp.1769-1777, 1997
- [19] M. McKinnon, G. Rouskas, H. Perros, "Performance Analysis of a Photonic Single-Hop ATM Switch Architecture with Tunable Transmitters and Fixed Frequency Receivers", *Performance Evaluation*, Vol.33, No.5, 1998
- [20] D. Gusfield, R.W. Irving, "The Stable Marriage Problem: Structure and Algorithms," MIT Press, 1989