Reconfigurable ATM Switch Fabrics Using Traffic History

H. S. Laskaridis, G. I. Papadimitriou, *Senior Member, IEEE*, and A. S. Pomportsis, *Member, IEEE*

***Abstract*—An ATM switch fabric which is capable of being reconfigured based on statistics of a previous time period, is introduced. Taking under consideration the strong correlation between ports in a campus or LAN ATM switch, the proposed architecture exhibits improved performance. We prove the performance improvement, by applying data collected from a campus production ATM switch onto our proposed architecture.**

***Index Terms*—ATM, correlation, reconfiguration, switch fabric.**

I. INTRODUCTION

THE rapid 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 gap between the capabilities of fiber and the capabilities of terminating, switching and routing equipment. 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, concerning buffering, packet synchronization, optical crosstalk, path-dependent loss [1]–[3]. These reasons make optical routing hard to implement. In a number of papers regarding photonic packet switches (e.g., [4], [5]), proposing per packet switching, O-E-O conversion is not avoided, although executed only for the headers. In [6]–[8] a number of open issues (e.g., contention resolution, header's extraction without O-E conversion), along with possible future solutions, are presented. Thus, given the present immaturity of optical technology, we consider O-E-O conversion mandatory.

In this letter, we propose a reconfigurable ATM switch fabric which is based on present-day optical and electronic technology. The proposed ATM switch fabric takes advantage of the strong correlation between ATM ports. A passive star coupler [9]–[11], is used in order for the switch to be easily reconfigured in the optical domain.

The rest of the letter is organized as follows. In Section II we demonstrate the correlation between ATM ports in the Aristotle University's network. In Section III we present the reconfigurable GASA architecture. In Section IV the performance improvement gained by applying reconfiguration is demonstrated using actual data. Finally, Section V concludes this article.

II. 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. Models using traffic distribution tables, with each input port having equal probability to transmit toward each output port, are far from realistic. In reality, there are pairs of input and output ports with no traffic. Thus, none of the cells that arrive at the specific input port is destined to the specific ouput port. This is merely due to the fact that 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 a one and only other end station.

The following graphs (see Fig. 1), created using Multi-Router Traffic Grapher (MRTG), reveal the "large time scale" correlation between two pairs of ATM ports from our university's campus network. Each graph shows the incoming (shadowed region) and outgoing (line) traffic (in megabits per second) of an ATM port for approximately 33 h.

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.

The idea of taking advantage of the traffic correlation, based on previously collected statistics, could be deployed even in routing protocols. For instance a modified PNNI protocol could make decisions regarding the establishment of new virtual circuits, based on the amount of anticipated traffic. However, in this letter we focus on designing an ATM switch which exploits the correlation, although the latter depends on the end stations that the user connects on the switch and the port that is used for each end station.

III. RECONFIGURABLE GASA ARCHITECTURE

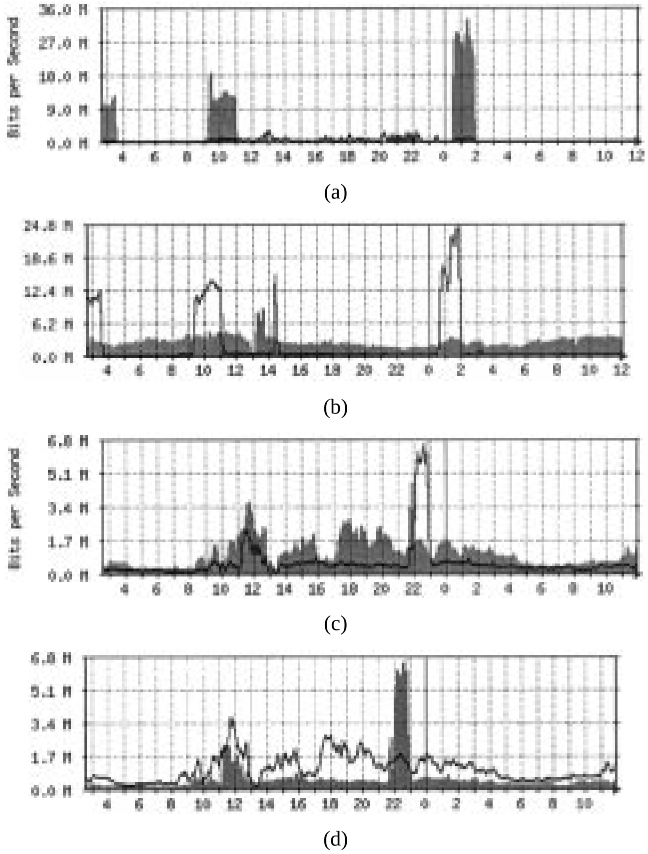
In [12] we have presented the grid-based ATM switch architecture (GASA). In a GASA switch an input port and an output port are connected on a switching element (SE). The SEs are interconnected forming a grid. Proper addressing of the SEs makes the implementation of a simple cell routing algorithm feasible.

In order to take advantage of the traffic correlation, we add an intermediate stage between input ports and SEs, prior to O/E conversion (Fig. 2). This stage consists of one tunable wavelength converter (TWC) for each input port, a passive star coupler and a fixed optical filter for each SE.

Manuscript received January 19, 2002. The associate editor coordinating the review of this letter and approving it for publication was Dr. J. Kim.

The authors are with the Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki 54006, Greece (e-mail: gp@csd.auth.gr).

Publisher Item Identifier 10.1109/LCOMM.2002.801341.



(a)

(b)

(c)

(d)

Fig. 1. Daily traffic statistics of various ATM ports.

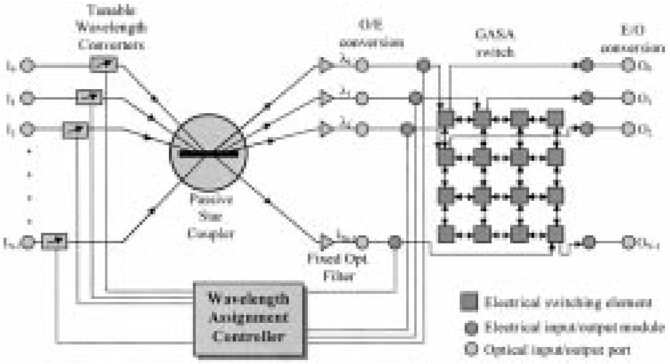


Fig. 2. Reconfigurable GASA.

Traffic statistics, regarding source and destination of each cell, are collected and transmitted to a control unit, namely, wavelength assignment controller (WAC). Periodically, WAC executes the “stable marriages” algorithm [13], 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 of each input and output port is easily extracted from the actual traffic matrix, which is created by collected measurements. The “stable marriages” algorithm maximizes the overall “satisfaction”.

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_z (through $\mathrm{SE_z}$ of course), WAC orders the TWC connected to I_x to transmit in wavelength λ_z. In this way optical signal coming from I_x will be received by the optical filter connected to SE_z. O_z is also connected to $\mathrm{SE_z}$, thus having high probability of switching the cell without crossing other SEs.

Thus the broadcast-and-select architecture is used to make the switch easily reconfigurable in the optical domain.

IV. PERFORMANCE IMPROVEMENT

The campus ATM network of Aristotle University of Thessaloniki (AUTh) consists of four 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: two 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 workstations and fast ethernet switches with ATM up-links.

We have collected data about the traffic switched through the border switch for two regular working days. We calculated the distribution of cells switched during the second day (approximately 2.5×10^{12} cells) based on the number of SEs (hops) that they would have to cross, in order to reach their destination, 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) previous half an hour; or c) corresponding half-hour period of the previous day.

Results are presented in Fig. 3. In case (1) we had an average number of hops 5.50; in case (2a) 4.17; in case (2b) 4.45; and in case (2c) 4.33. 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. This is because the reconfiguration algorithm attempts to 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. Thus reconfiguration based on the previous half hour performs better than other reconfiguration schemes, although differences are marginal.

We note that cell distributions presented in Fig. 3 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, keeping the order intact.

Note also that the frequency of executing the “stable marriages” algorithm (maximum twice per hour) is so low that cannot pause any performance restrictions, regarding the operation of WAC.

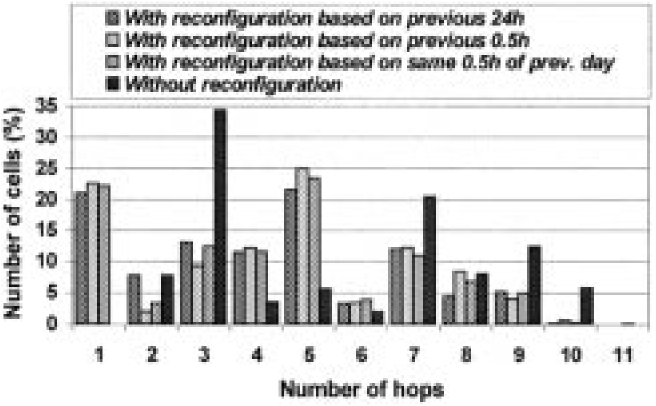


Fig. 3. Distribution of cells based on number of hops, with or without reconfiguration.

Exact traffic traces are not available due to technical reasons. Such traces, with exact arrival time on a per cell basis, would enable us to evaluate other performance metrics, such as delay and loss probability.

V. CONCLUSIONS AND FUTURE WORK

Increasing the number of cells being switched without crossing more than one SE intuitively 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 SEs.

Making the "output side" of the switch fabric also reconfigurable would further improve the performance. Arranging the connections of both input and output ports on the grid of SEs, in such a way that the average number of hops would be minimal, would result in further performance improvements. However, a much more sophisticated algorithm is required.

The idea of optical reconfiguration before entering the main switch fabric could also be applied to well-known architectures, such as Banyan networks, in order to decrease blocking probability and consequently, the length of necessary internal buffers. Of course, in this case a different wavelength assignment algorithm would be necessary.

REFERENCES

[1] L. Thylen, G. Karlsson, and O. Nilsson, "Switching technologies for future guided wave optical networks: Potentials and limitations of photonics and electronics," *IEEE Commun. Mag.*, vol. 34, no. 2, pp. 106–113, 1996.

[2] Y. Pan, C. Qiao, and Y. Yang, "Optical multistage interconnection networks: New challenges and approaches," *IEEE Commun. Mag.*, vol. 37, no. 2, pp. 50–56, 1999.

[3] B. Mukherjee, "WDM optical communication networks: Progress and challenges," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1810–1824, Oct. 2000.

[4] W. Zhong and R. Tucker, "Wavelength routing-based photonic packet buffers and their applications in photonic packet switching systems," *J. Lightwave Technol.*, vol. 16, pp. 1737–1745, Oct. 1998.

[5] J. Diao and P. Chu, "Analysis of partially shared buffering for WDM optical packet switching," *J. Lightwave Technol.*, vol. 17, pp. 2461–2469, Dec. 1999.

[6] S. Yao, B. Mukherjee, and S. Dixit, "Advances in photonic packet switching: An overview," *IEEE Commun. Mag.*, vol. 38, no. 2, pp. 84–94, 2000.

[7] D. Hunter and I. Andonovic, "Approaches to optical internet packet switching," *IEEE Commun. Mag.*, vol. 38, no. 9, pp. 116–122, 2000.

[8] K. Park and T. Mizumoto, "Packet header recognition assigning the position of a signal in the time axis and its application to all-optical self-routing," *J. Lightwave Technol.*, vol. 19, pp. 1076–1084, Aug. 2001.

[9] G. I. Papadimitriou and D. G. Maritsas, "Learning automata-based receiver conflict avoidance algorithms for WDM broadcast-and-select star networks," *IEEE/ACM Trans. Networking*, vol. 4, pp. 407–412, June 1996.

[10] G. I. Papadimitriou and A. S. Pomportsis, "Self-adaptive TDMA protocols for WDM star networks: A learning-automata-based approach," *IEEE Photon. Technol. Letters*, vol. 11, pp. 1322–1324, Oct. 1999.

[11] A. Misawa and M. Tsukada, "Broadcast-and-select photonic ATM switch with frequency division multiplexed output buffers," *J. Lightwave Technol.*, vol. 15, pp. 1769–1777, Oct. 1997.

[12] H. Laskaridis, A. Veglis, G. Papadimitriou, and A. Pomportsis, "Grid-based switch fabrics: A new approach in designing fault-tolerant ATM switches," *Comp. Commun.*, vol. 24, no. 15–16, pp. 1589–1606, 2001.

[13] D. Gusfield and R. W. Irving, *The Stable Marriage Problem: Structure and Algorithms*. Cambridge, MA: MIT Press, 1989.