TechLAN:
An Experimental High-Speed ATM LAN

Israel Cidon *Asad Khamisy* *Raphael Rom* *Yuval Shavitt* *Avraham Shmuel*
Moshe Sidi
Department of Electrical Engineering
Technion, Haifa, Israel

1 Introduction

ATM (Asynchronous Transfer Mode) is a set of international standards for sending large amounts of voice, data and video information simultaneously over a single network at speeds up to a thousand times faster than is commonplace today. It is aimed to transport virtually every electronic communications format from e-mail to phone calls to movies. ATM promises to be a true worldwide standard and the foundation for the next century networks.

The most common ATM network configuration is based on one or more fast switches interconnecting the subscribers. Such switches render the entire network distance independent, namely, the same bandwidth and service can be afforded to the communicating partners whether 1 meter or 100 kilometers apart. Unfortunately, this independence comes at a fairly high cost of the switch. It appears that when the communication parties span only a small locale, different configurations can be devised that rid themselves from reliance on the costly switch. TechLAN is such a network and is being developed at the Computer Networks Laboratory in the Department of Electrical Engineering at the Technion. Its purpose is twofold; to provide a low cost solution to share an ATM switch connection among several ATM users, and to serve as a LAN which provides local interconnection between ATM users without interfering with the traffic of the ATM switch connection. The combination of these two goals, and especially the need to support Quality of Service (QOS) parameters associated with ATM connections within shared environment is particularly challenging. TechLAN achieves these two goals using Buffer Insertion Ring and Shared Bus Switch architectures to be described later.

TechLAN will provide the foundation over which new networking techniques and new applications will be experimented with. On this experimental network we plan to develop and study various applications, such as those related to teaching and technological education. Another important activity is to study the behavior and performance of the network under various traffic types such as voice, video and data, and under different load conditions.

2 Network Structure

The network is made up of rings that are connected with high speed interconnection devices (termed miniswitches). The general structure is depicted in Figure 1. This modular structure allows more miniswitches and rings to be added easily. Network interface cards (NIC) are available for PCs and are under development for SUN workstations. Miniswitch cards are PC-based. Several application cards, including multimedia cards, cell generator cards, network analyzer cards and others, that can be connected to the NIC were implemented. In the following we describe the various components of the network.

The ring is built using buffer insertion ring (BIR) architecture [3, 7] with cell removal by the destination [6]. Such a ring does not use tokens for its transmission scheduling and allows spatial reuse, i.e., it can support simultaneous transmissions over different segments of the network. As a result, the total ring throughput can be significantly higher than the capacity of a single link (which limits other ring architectures) . The principle of the BIR operation is as follows [3]. On the receiving side of each link, there is an insertion buffer (IB), which can store one ATM cell. A node may start a cell transmission at any time if its IB is empty and if there is no ring traffic on its incoming link. If ring traffic is arriving while the node is in the middle of a cell transmission, then this traffic is stored in the

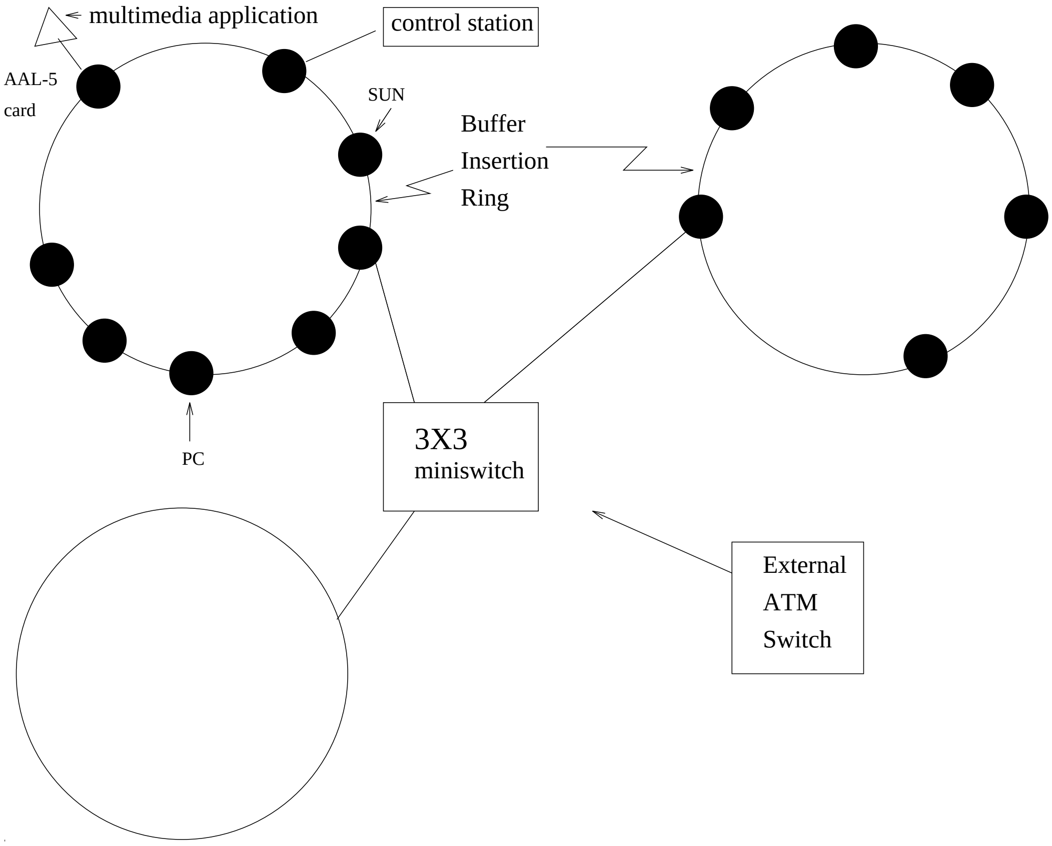

Figure 1: General network structure

IB until the cell transmission is complete. The node cannot transmit any more until the IB becomes empty and there is no ring traffic on the link. In other words, non-preemptive priority is given to the ring traffic. If a node is idle, ring traffic not destined to that node cuts through the IB, without delay.

Since ring traffic has priority, it is clear that without an additional access control mechanism heavily loaded nodes can prevent other nodes from accessing the ring – a situation known as “starvation”. To prevent this problem, mechanisms that guarantee fair access to all nodes have been proposed for BIRs [5, 6, 4]. Some of these mechanisms are used in TechLAN together with a bandwidth reservation scheme as required by ATM standards.

The NIC can receive (transmit) input (output) from (to) two sources (see figure 2): the local bus of its host computer (ISA for PC, SBUS for SUN workstations) and an auxiliary connector that can be connected directly to four I/O devices such as multimedia cards. The direct connection to the host via its bus is implemented using buffering FIFOs (part of the buffering unit of figure 2) due to the low bandwidth of the bus. The auxiliary connection, on the other hand, is not buffered, requiring the connected units to receive data at the network speed or provide its own buffering.

The NICs are physically connected to one another via differential coax cables that are driven by AMD TAXI chips [2] at data rate of 100Mbps. The TAXI chips can also drive fiber optics; the selection of coax cable was made solely due to cost consideration. Under development is a newer version of the NICs that is based on Cypress HOTLink chips [1] and is using fiber optics at rates around 200Mbps. Both the TAXI and the HOTLink chips encode 3/4/5 bits of data to 4/5/6 transmitted bits to make communication more immune to noise. This transformation enables the chip to support in-band control characters which are used by the ring medium access protocol for signaling.

When a cell arrives at the receive unit (figure 2) its VCI is checked to identify if it should be removed from the ring and what is its destination. The cell destination can be any combination of the host computer and the four auxiliary port devices. The ability

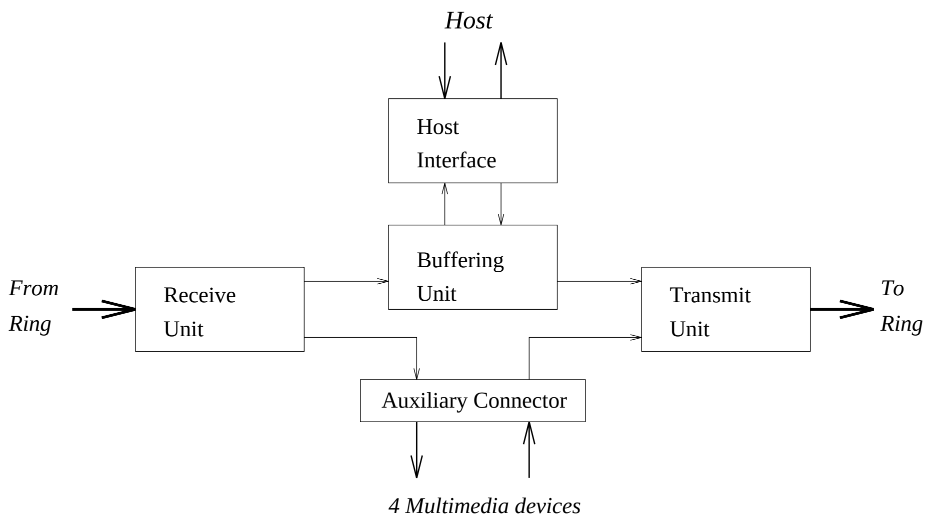

Figure 2: NIC general structure

to copy and repeat cells simultaneously supports efficient multicast implementation. The VCI table is programmed by the NIC host and is, currently, 12-bit deep. The receive unit also delivers the in-band control signals to a control hardware (not shown in figure 2) that implements the MAC algorithms.

The buffering unit contains the IB and the FIFOs that store data destined to the host and coming from the host. The transmit unit can transmit (in decreasing priority order) an in-band control signal, a cell from the IB, a cell from the auxiliary port, or a cell from the host buffer.

The interconnection device (miniswitch) is a PC card that can connect three rings as well as its local host. It is built using a shared bus architecture with output buffering [8]. This architecture was selected due to its simplicity and because interconnecting a larger number of rings is not anticipated. In addition, the miniswitch supports multicast transmission which is important for multimedia conferencing applications. The miniswitch structure consists of four I/O ports, a pipeline unit (which implements a fast and wide bus) and a single RAM-based routing table shared by all four ports of the miniswitch. The pipeline unit performs the ATM label swapping, and checks the ATM HEC to identify and discard cells with faulty headers. The miniswitch card can reside within a PC station which enables the implementation of complicated network control mechanisms.

2.1 Software and Control

The software is organized in several layers: the ATM layer that also includes a control and management plane, the AAL5 layer, and an IP layer on top of AAL5. The ATM layer is responsible for transferring cells to and from the NIC and performs flow control functions such as the GCRA rate control mechanism of the ATM. Standard ATM UNI signaling is supported. The AAL5 layer implements segmentation and reassembly in message-mode only (in addition to the AAL5 header generation). The IP layer itself is divided into two sublayers, the lower one contains also the ARP protocol and is adapted to handle the ATM connection oriented communication.

One of the stations in every ring is elected as a *control station* and is responsible for managing connection set-up and take-down as well as bandwidth allocation to connections. This mechanism is implemented in the ATM layer. Setting up connections between stations on different rings is done by the control stations and the hosts that control the switches enroute. Error recovery functions such as control station failing, connection management and other functions are also implemented in the management plane.

External connections (to hosts outside the set of interconnected rings) can be done in two distinct ways. Since IP is supported by the software, any host with TechLAN card and a connection to the Internet can be easily configured as a gateway. Connection to an external (public) ATM switch will be implemented by installing an ATM UNI card in the switch; preliminary software support for such a connection is being designed.

3 Discussion

Putting the emphasis on simplicity, we have built a cheap high-speed ATM LAN. In addition to the NIC and the switch described above we have built application cards that can be connected to the NIC auxiliary port: mainly multimedia cards and performance analysis cards.

The multimedia cards we have in operation are audio/video/data generator and a multi-window browser. In the currently available cards the video information is not compressed. A second version of the set of audio/video generator/receiver that use compressed video are currently being built.

Performance analysis of the network is done using two cards: a cell-based traffic generator card which can generate ATM connections with different parameters (sustained rate, peak rate and burst size) and different durations; and a network analyzer that can collect statistics about the streams it monitors.

Acknowledgment

Many of our students have contributed to the design and development of this network, particularly Noam Anchel and Shy Samoha. We acknowledge the generous support of SUN Microsystems, Tadiran Tikshoret, Intel, Rad and Zoran which enabled to conduct this development.

References

[1] *CYPRESS CY7B923/933 HOTLink transmitter/receiver.*

[2] *AMD Am7968/Am7969 TAXIchip*, 1992.

[3] D. Bertsekas and R. Gallager. *Data Networks.* Prentice Hall, New-Jersey, 1987.

[4] Israel Cidon, Leonidas Georgiadis, Roch Guérin, and Yuval Shavitt. Improving delays and throughput of fair algorithms for networks with spatial reuse. In *INFOCOM'94*, pages 1103 –1111, June 1994.

[5] Israel Cidon and Yoram Ofek. Distributed fairness algorithms for local area networks with concurrent transmissions. In *3rd International Workshop on Distributed Algorithms*, pages 57 – 69. Springer-Verlag, September 1989. Lecture Notes in Computer Science 392.

[6] Israel Cidon and Yoram Ofek. Metaring – a full duplex ring with fairness and spatial reuse. *IEEE Trans. Commun.*, 41(1):110–120, January 1993. (See also Proc. INFOCOM'90).

[7] S. Leventis, G. Papadopolous, S. Koubias, and J. Constantinides. A protocol for a new double-loop computer network and its implementation. In *Proc. of the National Computer Conference (NCC)*, pages 203–208, Chicago IL., May 1981.

[8] F. A. Tobagi. Fast packet switch architectures for broadband integrated services digital networks. *Proceedings of the IEEE*, 78(1):133–167, January 1990.