Reprint

The Physical Layer of a Cell-based Local Area Network

A. Maniatopoulos, T. Antonakopoulos and V. Makios

The 4th International Conference on Advances in Communication and Control – COMCON'4

RHODES, GREECE, JUNE 1993

The Physical Layer of a Cell-based Local Area Network

A. MANIATOPOULOS, T. ANTONAKOPOULOS AND V. MAKIOS
Laboratory of Electromagnetics, Electrical Engineering Department
University of Patras, 26110 Patras, Greece

Abstract-- The functional description and the implementation of an experimental research prototype of the physical layer of a cell-based high-speed local area network is described. The prototype operates at 800 Mbits/sec and is based on the functionality resulting from combining functions of the Fibre Channel with the basic features of the pure ATM of the B-ISDN protocol reference model. Although most of the physical layer functions have been implemented using Field Programable Gate Arrays, high cell processing rate (1.88Mcells/sec) has been achieved by using a parallel architecture in the various processing stages. The developed physical layer could be used in local area networks, in customer premises networks or for the user-network interface in a broadband network.

I. INTRODUCTION

It is unquestionable that the two basic requirements of new developments in the Local Area Networks (LAN) environment for supporting various types of traffic efficiently, are the service-independence and the high transmission rate. The most promising technique for achieving the service-independent goal is the Asynchronous Transfer Mode (ATM) used in the Broadband ISDN (B-ISDN) [1], [2]. The ATM uses short, fixed length cells for user information transmission. The cells are switched using the routing information of their header and various types of traffic are effectively supported by using adaptation functions. In this paper the physical layer of a Buffer Insertion Cell-based LAN (BIC-LAN) is described.

The used access method is applicable in ring topologies for supporting various types of traffic using cells in a slotted operation with 'destination release' [3]. The method incorporates a distributed algorithm for allocating the bandwidth to the various stations by comparing the total requested bandwidth for synchronous traffic to the actually measured traffic and adjusting the offered load accordingly. The access method ensures the transmission of synchronous traffic by fulfilling its Quality of Service (QOS) requirements in terms of transmission delay variations and cell rejection rate.

The Physical Layer has to provide bit and byte timing synchronization, asynchronous cell multiplexing in the physical medium and header error detection and correction. The network stations are connected using point-to-point links and during each transmission, constant delay is imposed by each intermediate station. This delay depends on the transmission speed and the cell length. The developed Physical layer is based on the functionality of the Fibre Channel [4] and the B-ISDN protocol reference model [2].

Section II gives a concise description of the used medium access protocol while the architecture of a network station is given in section III where the basic requirements of the ATM layer from the Physical layer are determined. In Section IV the Physical layer structure is described in details, emphasizing on how cell processing is performed in high transmission speeds and how the Fibre Channel technology is used in a cell-based environment.

II. THE MEDIUM ACCESS PROTOCOL

The BIC access method has been based on the buffer insertion technique. Each station receives the upstream traffic, rejects the idle and error cells, removes the cells destined to this station and transmits its own cells if there are no upstream cells in its ring buffer. Using various network statistics the access method determines if a station cell will be transmitted or the station must defer. The multiplexing of the various station services is performed by the access method using the QOS parameters determined during each service connection establishment. The idle cells are used to insert new data in the cell stream or for decoupling the differences in the transmission speed between adjacent nodes. The basic advantages of the BIC access method are its ability to allocate a portion of the available bandwidth to a specific service by using a distributed algorithm under the use of network statistics to keep the total offered load a target value. This is achieved by controlling the flow of synchronous services (connection-oriented) using the initially negotiated values and by reducing the asynchronous traffic available bandwidth. In order to minimize the effect of the network traffic to the delay variations of the same synchronous stream, a traffic monitoring method is used to control the type and the amount of the offered load.

The offered-load control algorithm used in this access method is based on traffic monitoring and is fully distributed since each station decides on its offered load independently from the other stations by using its traffic monitoring statistics. This control algorithm uses the basic concept of LOCOST [5] which states that the appropriate scheduling of traffic can improve the network performance and if it is properly managed, it can change the characteristics of the access method.

The basic parameter of the access method is the Total Traffic Window which gives the latest part of the traffic used by the algorithm to collect its statistics. The 'Total Traffic Window' is an indirect expression of the available bandwidth and is divided in two parts: the first part is devoted exclusively to the synchronous traffic and the second part is used to support both classes of traffic. The portion of the bandwidth devoted to the synchronous traffic is determined by the Connections' Window and the instantaneous synchronous traffic measurement. The Connections' Window is the sum of the number of cells generated during the Total Traffic Window by each synchronous connection. It is updated each time a new connection is established or when the connection is released. The instantaneous synchronous traffic measurement indicates the number of synchronous cells measured during the Total Traffic Window.

When a new synchronous connection must be established, the station estimates the mean cell rate generated by the specific service and if there is available bandwidth the connection is accepted and a Connections' Window Update procedure begins. Otherwise, the connection is rejected since there is no available bandwidth to support the specific service effectively. During the Connections' Window Update procedure, the node transmits a control cell with its own address as the destination address. This control cell passes through all of the network

nodes, so each node is informed about the new status of the synchronous traffic connections. The same procedure is used when a connection is released and the Connections' Window Update procedure informs all nodes to decrease their Connection Window appropriately. Using this mechanism, all stations utilize the same protocol parameters and the protocol fairness is guaranteed. When the node has no synchronous cells to transmit and it detects that there is no cell in its ring buffer, the transmission of an asynchronous cell depends on the bandwidth percentage allocated exclusively for synchronous traffic. The smaller the bandwidth portion allocated to synchronous traffic, the higher the probability of asynchronous transmission.

The access method gives the requested bandwidth to the synchronous traffic by varying the asynchronous offered load. The synchronous traffic throughput of a node varies around its mean value for a given number of connections, depending on the total statistical multiplexing of the synchronous traffic. When the portion of the transmitted synchronous cells to the total traffic decreases, the algorithm estimates that more bandwidth is available for asynchronous traffic and the asynchronous offered load increases.

III. THE NETWORK STATION STRUCTURE

The block diagram of the station architecture is shown in Fig. 1. The ATM layer performs the BIC access protocol functions and provides the multiplexing/demultiplexing of the cells of various connections into the network cell stream. The ATM receives the payload of each cell from the respective service adaptation unit and transmits it into the network by adding its header. The cell headers are constructed using the information provided by the Station Management Unit (SMU). The ATM layer interfaces with the SMU and the various service adaptation units using an Internal Bus which is composed of two high speed (100 Mbytes/sec) data transfer paths (one for cell transmission and the other for cell reception) and a control bus. The control bus operates under the supervision of SMU. The ATM layer contains two 32-bit wide FIFOs for temporary cell payload storage during cell header processing in the ATM layer. The 'Tx-Rx Multiplexer' interconnects the two FIFOs when the ATM Controller detects that the cell just received from the network has to be retransmitted into the network. The ATM layer provides also the link to the SMU for monitoring the physical layer functions. In each direction the ATM layer uses two parallel buses for exchanging information with the Physical layer. A 32-bit bus is used for the cell payload and an 8-bit bus for the cell header. In the transmit direction the header of a cell is transferred to the Physical layer during the transmission of the payload of the previous cell. In the receive direction, the cell header is transmitted from the Physical layer to the ATM layer simultaneously with the first four bytes of the cell payload.

A part of the Physical Layer is based on the Fibre Channel architecture. It uses the FC-0 functions of the Fibre Channel (transmission media adaptation, transmitters, receivers and their interfaces) and a combination of the 8B/10B and 4B1C transmission codes. The Physical layer uses some of the FC-1 functions, bit and Transmission-Word boundaries synchronization and also contains the required functionality for adaptation of the ATM cells (idle and information) to the Fibre Channel Datagrams. As a Fibre Channel Datagram is considered the information (number of cells) contained between two idle cells. Each idle cell is used simultaneously as the End Delimiter of the current datagram and as the Start Delimiter of the new datagram. The length of a datagram may vary between zero (two consecutive idle cells) and a maximum number of cells determined by the used cell rate decoupling parameter. The Datagrams are reconstructed when they pass through a network station and new frame delimiters are generated while the contained data (cell

burst) are modified. This modification may include idle cell substitution with station data cells or the insertion of idle cells for supplementing the used 'destination release' scheme at the ATM layer.

In the transmit direction the Physical layer accepts complete cells from the ATM and generates the fifth byte (Header Check Sequence-HEC) of the cell header. Then the cell is passed through the Fibre Channel Framer module which generates the appropriate FC frame format and multiplexes the ATM cells in its structure. The FC Framer uses a TAXI transmitter for interfacing the physical medium. At the receiving side, a TAXI receiver accepts encoded serial data and generates encoded parallel data to the receive part of the FC Framer. The FC Framer recognizes the data semantics, removes the FC delimiters, regenerates the ATM cells and validates the HEC field of their header. Single-bit error correction is performed at the header and the cells are passed to the ATM layer for further processing.

IV. THE PHYSICAL LAYER STRUCTURE

The Physical layer is subdivided into two sublayers. The lowest sublayer, which is called Fibre Channel Adapter (FCA), is related with the functionality of the Fibre Channel and deals with the physical medium aspects. The higher sublayer, which is called Transmission Convergence (TC) sublayer, deals with the cell stream multiplexing/demultiplexing, cell header error detection and correction, idle cell insertion and extraction and coding/decoding of data.

The data are encoded using a combination of the 8B/10B [4] and the B1C [6] codes. According to the FC-1 transmission code, information transmitted over FC shall be encoded eight bits at a time into a 10-bit transmission character and sent serially. At the receiver the encoded character-must be decoded back to 8-bits. The problem of using the 8B/10B code in the ATM environment with single error correction is that due to the coding, an error bit in the transmitted serial stream may be spread into two bits in the decoded data, making the single error correction capability of the Transmission Convergence sublayer useless. For that reason, the user cells are encoded using the 4B1C code while the idle cells, which are used at the physical layer for cell rate decoupling and cell delineation, are constructed as a sequense of K28.5 (FC terminology) bytes and its violation for conformance with the 4B1C code. Following the above mentioned coding scheme the Fibre Channel requirement that the maximum run-length of the encoded data does not exceed five bits is satisfied and the correct phase relationship of the extracted clock with the received data is maintained.

Since in the BIC-LAN fixed length cells are used a cell delineation mechanism is implemented. This mechanism is based on the 'Header Error Check" method and on the 'Use of emply cells'. The idle cells in the BIC-LAN are used for cell rate decoupling between adjacent stations at the physical layer and for bandwidth allocation at the ATM layer. The idle cells are generated in the Transmission Convergence sublayer when there is no cell available from the ATM layer or when the number of user cells transmitted after the last idle cell has exceeded a predetermined threshold. At the receiving side, the idle cells are detected and removed at the TC sublayer where the cell delineation mechanism is applied.

The idle cells contain four K28.5 bytes in the first four bytes of their header, while the idle cell payload contains twelve four-byte blocks where the first byte of each block is a K28.5 character and the other three are corrected K28.5 (CK 28.5) characters in order to follow the 4B1C rules (the sixth bit of CK 28.5 is the complementary of the seventh bit instead of being equal to it as in the K28.5). Acccording to the general cell delineation state diagram of [1], in the HUNT state, the

synchronizer checks for four consecutive K28.5 bytes. Bit synchronization is achieved at the Fibre Channel Adapter. In this case, the synchronizer changes to the PRESYNC state and checks the idle cell payload for the pattern of twelve (K28.5) (CK 28.5) (CK 28.5) (CK 28.5) bytes. When the cell boundary is confirmed, the synchronizer switches to SYNC state. When m consecutive cells have been received with an error indicated during the validation of their HEC, the system is considered to be desynchronized and a new idle cell is needed for resynchronization. The above mentioned synchronization mechanism has the same performance with the mechanism used in B-ISDN during loss of synchronization while, during normal system operation, it has a faster synchronization phase.

A. The TC Sublayer Transmitter

The TC transmitter shown in Fig. 2 is used to receive user cells from the ATM layer, to generate the HEC byte of the cells' header and to generate the required idle cells for cell rate decoupling and bandwidth allocation, since in the BIC-LAN there are no unassigned cells for use in the ATM layer. The TC interface to the ATM includes the P-Bus (B1:0) for cell payload and the H-bus (7:0) for cell header transfer. The transfer speed at the P-Bus is 100Mbytes/sec while in the H-bus is 25 Mbytes/sec. The data at the two buses are processed independently and are multiplexed with the output of the Idle Cell Generator in the multiplexing stage of the Tx FC Frame. As it is shown in Fig. 3, after the transmission of the fortieth byte of the payload of a cell (NCI pulse) the TC transmitter checks the availability of a new cell in the ATM layer (IDLE signal). If no cell is available the Idle Cell Generator is activated. In the case where a new cell is available, the first four bytes of the cell header are transferred and stored in the Pipeline module of the FC Frames while the HEC byte is calculated and the 4B1C coding is performed. When the payload transmission of the previous cell has been finished, the header of the next cell is stored in the Physical layer pipeline structure and is ready for transmission. Although the operational need of the circuit is 100MHz, this parallel architecture allows the CRC generation to be performed at 25MHz. Since the TC transmitter has been implemented in a XC3164 FPGA and in order to achieve the total of 100 Mbytes/sec system operation, the 32-bit wide P-Bus is multiplexed in three stages and the 4B1C coding circuit has been implemented in four different transfer paths.

B. The TC Sublayer Receiver

The TC receiver which is shown in Fig. 4 performs the receive functions of the physical layer, such as cell delineation, HEC validation and correction, idle cell extraction etc. The FCA modules tranfer to the TC receiver the encoded data and an indication that a K28.5 pattern has been detected. The RESYNC indication along with the output of the Idle Cell detector and the HEC Error Detection signal are the required information for performing the previously mentioned synchronization method in the Cell Synchronization module of the Rx FC Frame. As it is shown in Fig. 5, the first four bytes of the received header are decoded and stored in the Head Buffer for eight byte cycles. The payload bytes received during this period are also stored in the Payload Buffer and the possible HEC correction is performed. If more than a single error have been detected in the received header, the cell is disgarded, otherwise the position (byte number and bit position) of the error is calculated and the header correction is performed during the transfer of the header to the ATM layer using a XOR-gate plane. The TC receiver has also been implemented using a XC3164 FPGA.

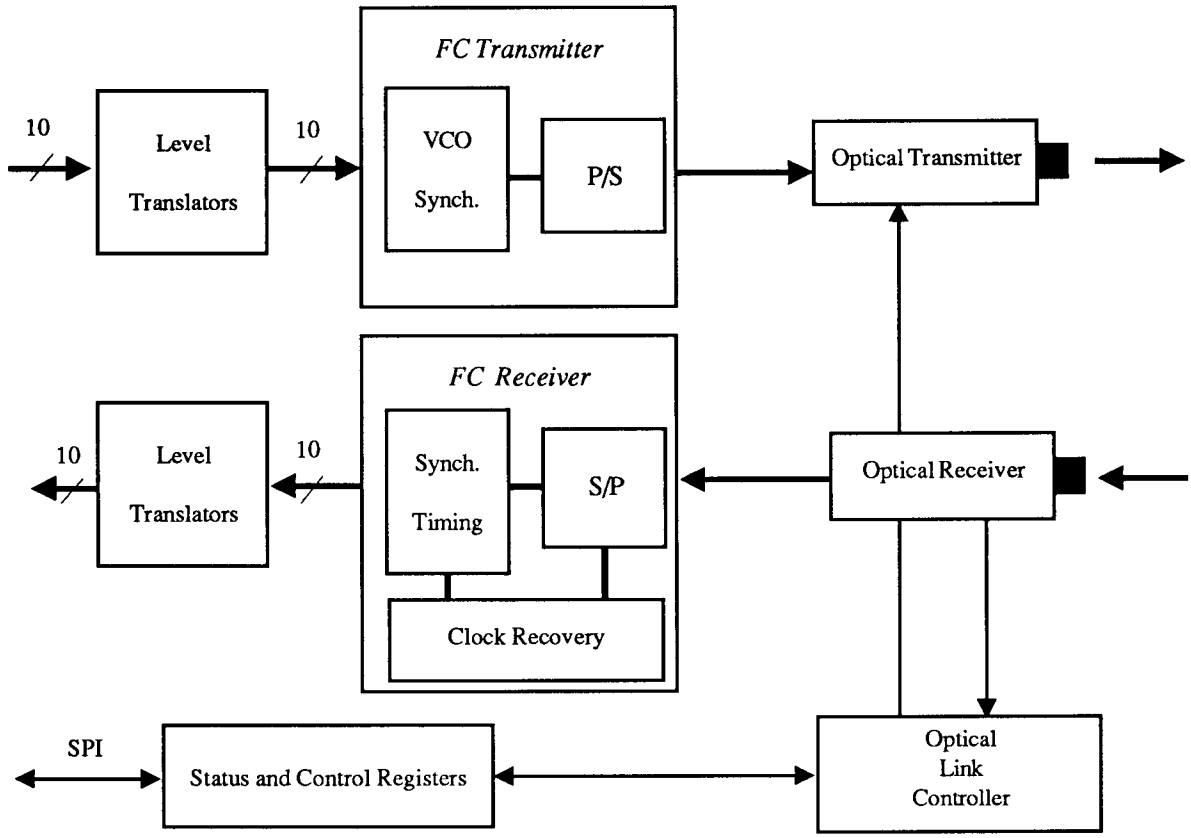

Fig. 6. Fibre Channel Adapter block diagram.

C. The Fiber Channel Adapter

The Fibre Channel Adapter receives TTL-level signals and operates in PECL levels. For this reason, in the interface between TC and FCA TTL/PECL and PECL/TTL level translators have been used. The FC Transmitter (Fig. 6) is a commercially available TAXI IC which receives a 10-bit parallel bus and generates the NRZ bit stream for driving the optical transmitter. The FC Receiver accepts the NRZ serial data from the optical receiver, recovers the data clock by using its internal clock recovery PLL, re-establishes byte boundaries using the SYNC (K28.5) pattern and generates 10-bit synchronous parallel bus for further processing by the Transmission Convergence sublayer.

For the implementation of the optical interface, the FTM-8500 (Tx) and FRM-8500 (Rx) data link modules of FINISAR Corp. have been used. These modules use short wavelength lasers to transmit data at Gb/s rates through multi-mode fiber (62.5/125). They have been operated in 1 Gb/s with 10^{-14} bit error rate. These two optical modules are controlled by the FCC-2000 controller achieving full control and performance monitoring of the upstream and downstream optical links. In the transmitting side the optical power launched into the fiber is controlled while the received optical power in the photodiode is measured in combination with the module temperature.

These facilities allow the Station's Management Unit to determine the status of its upstream and downstream links, allowing the higher layers' management to take the appropriate corrective actions during a failure in the optical link. In dual ring architectures, the optical link performance monitoring along with the loopback capability of the TAXI components can be used for fault-tolerant system operation, since the detection of an optical failure forces the TAXI Tx and Rx in the loopback mode and the network operates as a single ring as long as the failure is present.

VI. CONCLUSIONS

In this paper the detailed design and implementation of the physical layer of a cell-based LAN has been presented. The presented experimental prototype performs the required functions (cell delineation, cell header error detection and correction, synchronization etc) for transmitting a cell stream using the basic functions of the Fibre Channel. The required high cell processing rate is achieved by using a parallel architecture in the Transmission Convergence sublayer having different paths for the cell header and the cell payload and multiplexing these paths near the FCA interface. The used coding and cell delineation methods allow fast synchronization and single-error correction in the cell header. The cell processing capability of the developed prototype is 1.88Mcells/sec. The Transmitter and Receiver have been implemented using Field Programmable Gate Arrays (FPGAs), while the Fibre Channel Adapter transmit and receive modules have been implemented using commercially available ICs.

REFERENCES

[1] Martin de Prycker: "Asynchronous Transfer Mode: Solution for Broadband ISDN", Ellis Horwood Ltd, 1991.

[2] Michel Devault, Jean-Yves Cochennec, and Michel Servel, "The "Prelude" ATD Experiment: Assessments and Future Prospects", *IEEE Journal on Selected Areas in Communications*, Vol. 6, No. 9, December 1988, pp. 1528-1536.

[3] T. Antonakopoulos, S. Koutroubinas, J. Koutsonikos and V. Makios: " A Distributed Bandwidth Allocation Algorithm for Gbit/s LANs" to be presented in *the 19th EUROMICRO Conference*, Barcelona, Spain, September 1993.

[4] "Fibre Channel - Physical and Signaling Interface (FC-PH) Rev. 3.0", ANSI X3T9 I/O Interface, *Working Draft*, June 1992.

[5] John O. Limb: "Load-Controlled Scheduling of Traffic on High-Speed Metropolitan Area Networks", IEEE Trans. on Communications, Vol. 37, No 11, November 1989, pp. 1144-1150.

[6] Noriaki Yoshikai, Koh-Ichi Katagiri and Takeshi Ito: "mB1C Code and Its Performance in an Optical Communication System", *IEEE Transactions on Communications*, Vol. Com-32, No. 2, February 1984, pp. 163-168.