IMPLEMENTATION OF GIGABIT ETHERNET STANDARD USING FPGA

V.R. Gad[1] ,R. S. Gad[1] and G.M. Naik[1]

[1]Department of Electronics, Goa University, Goa, India
vinaya_gad@rediffmail.com

ABSTRACT

This paper presents the results of Gigabit Ethernet standard implementation in the FPGA device. The design uses Altera' s Stratix-II GX device and supports data transfer rates of 10Mbps, 100Mpbs as well as 1Gbps. FPGA implementations have the advantage of altering the functionality of the platform to perform several tasks. The performance of this design is assessed for Gigabit Ethernet using optical fibre as well as Copper media.

KEYWORDS

Gigabit Ethernet, FPGA, SFP, Frame length,

1. INTRODUCTION

The Gigabit Ethernet technology is an extension of the 10/100-Mbps Ethernet standard. Gigabit Ethernet provides a raw data bandwidth of 1000 Mbps while maintaining full compatibility with the installed base of over 70 million Ethernet nodes [1]. Gigabit Ethernet includes both full- and half-duplex operating modes. A Gigabit Ethernet is imperative for two reasons: faster systems and faster backbones. Gigabit Ethernet has the potential for low-cost products, freedom of choice in selecting the products, interoperability, and backward compatibility. Gigabit Ethernet supports existing applications, network operating systems, and network management; it requires a minimal learning curve for Ethernet network administrators and users. These investment preservation and risk minimization aspects are what make Gigabit Ethernet so attractive. With the development of Ethernet systems and the growing capacity of modern silicon technology, embedded communication networks are playing an increasingly important role in embedded and safety critical systems [2]. Advances in VLSI technology have also pushed integration to the point where it is now possible to design and implement a microprocessor and network controller on a single chip, known as System-on-Chip (SoC). In a network of embedded systems, each system can communicate with the other systems in the network, sharing information and sending and responding to requests as needed. Embedded devices need to be designed to solve specific problems. It is a challenge to find the right balance between power and cost. This becomes even more complicated when adding network capability to a device. The advent of Field Programmable Gate Arrays (FPGAs) with thousands of logic gates has made it possible to verify specific software functions on specific hardware. This reduces the design cycle and hence the execution cycle time to make the embedded system respond faster in real-time. Gigabit Ethernet applications which require line-rate processing for all frame sizes require hardware implementation [3]. A reconfigurable NIC (Network Interface Card) allows rapid prototyping of

DOI : 10.5121/ijmnct.2012.2404

new system architectures for network interfaces. The architectures can be verified in real environment, and potential implementation bottlenecks can be identified [4]. Dynamically reconfigurable platform will also reduce power consumption of the network device [5].FPGA based-platforms fulfill the performance requirements and provide extra flexibility in comparison to ASIC implementations. Using an FPGA based custom design PCI platform, we have incorporated a 10/100/1000Mbps MAC design to build a low cost, high performance embedded network controller. The only hardware required is a physical network interface, provided in this case by a standard Small Form-factor Pluggable (SFP) module. Figure 1 and Figure 2 gives an overview of the IEEE 802.3 Ethernet Standards.

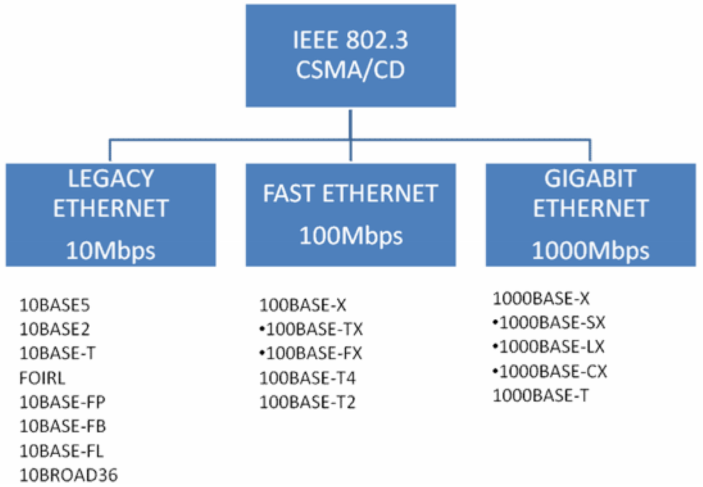


Figure 1.IEEE 802.3 Ethernet Standards overview

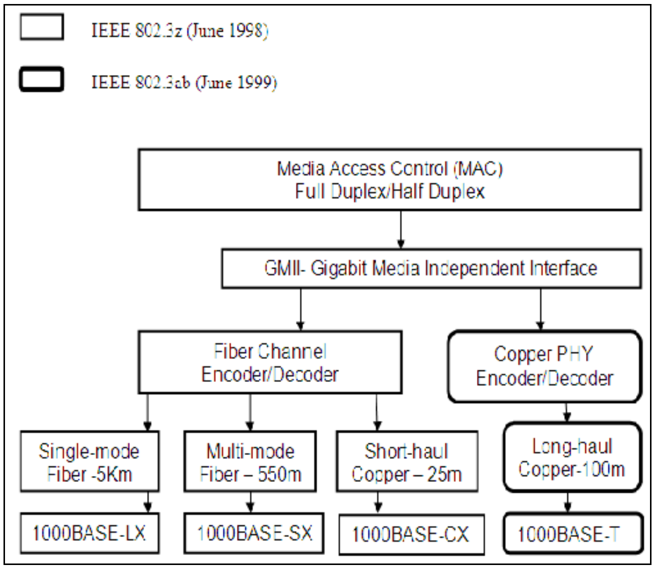


Figure 2.Gigabit Ethernet Standard [6]

2. Ethernet Media Access Layer Implementation

The heart of the Ethernet system is the frame. The network hardware comprising of the Ethernet interfaces, media cables move Ethernet frames between computers or stations. The bits in the Ethernet frame are formed up in specified fields. Figure 3 gives the basic Ethernet Frame Format.

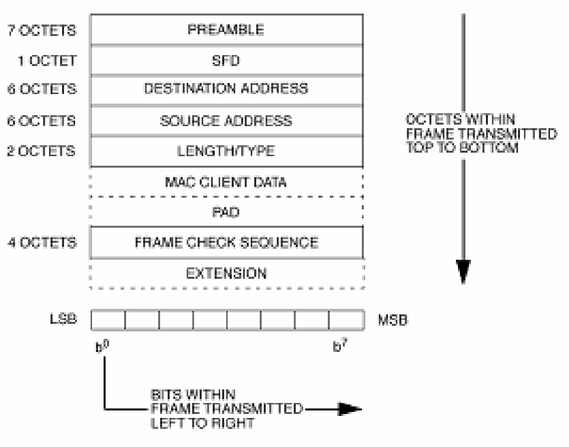


Figure 3. Basic Ethernet Frame Format [7]

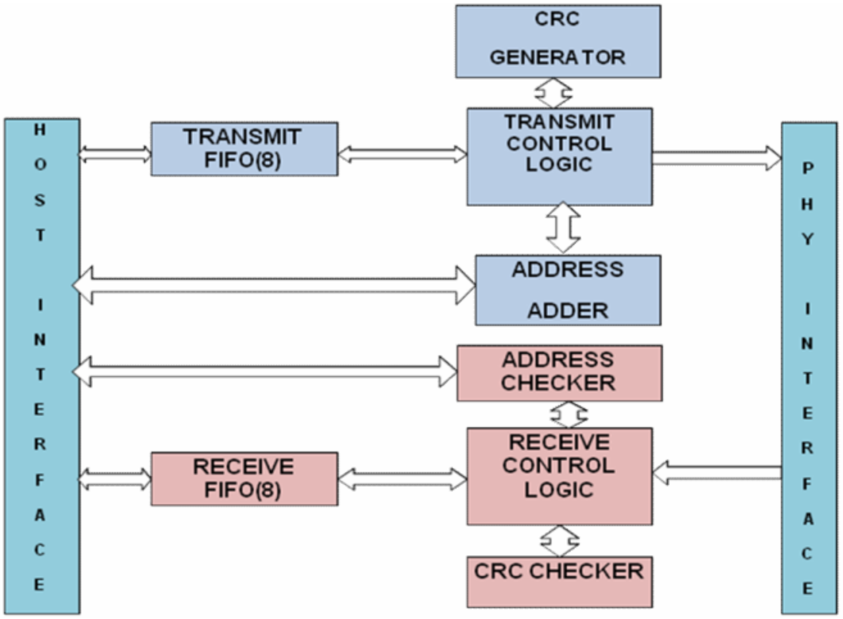


Figure 4. MAC Implementation Block Diagram

The Ethernet MAC (Media Access Control) architecture shown in Figure 4 consists mainly of the Transmitter and the Receiver.

- For the transmitter there is a Transmitter FIFO (First In First Out) (8 bit, 1500 bytes) where the data received from host is stored before it is given to the transmitter control module.
- The receiver has an Receiver FIFO (8 bit, 1504 bytes) where it stores the data which it receives from the receiver control module.

From the host interface all the signals are generated which are needed by the MAC module and the respective FIFO buffers.

- The transmitter part contains the Transmitter control module which generates frames.
- The MAC address adder adds the destination address, the source address and the type/length field of the frame which is generated in the transmit control module.
- A CRC (Cyclic Redundancy Code) generator is included in the transmitter to generate the required CRC bits which is appended as the FCS(Frame Check Sequence).
- The receiver part contains the Receiver control module which decapsulates the frame.
- A CRC checker is used to check for errors in transmission.
- MAC address checker is used for MAC address checking. It also checks for multicast and broadcast packets.

Figure 5 illustrates the programmable 10/100/1000Mbps Ethernet operation.The 10/100/1000 Ethernet PHY devices implement a shared interface that you connect to a 10/100-Mbps MAC via MII/RGMII (Reduced Gigabit Media Independent Interface) or to a gigabit MAC via GMII/RGMII [8]. On the receive path, the clock provided by the PHY device (2.5 MHz, 25 MHz or 125 MHz) is connected to the MAC clock, rx_clk. The PHY interface is connected to both the MII (active PHY signals) and GMII of the MAC function. On the transmit path, standard programmable PHY devices operating in 10/100 mode generate a 2.5 MHz (10 Mbps) or a 25 MHz (100 Mbps) clock. In gigabit mode, the PHY device expects a 125-MHz clock from the MAC function. Because the MAC function does not generate a clock output, an external clock module is introduced to drive the 125 MHz clock to the MAC function and PHY devices. In 10/100 mode, the clock generated by the MAC to the PHY can be tri-stated. During transmission, the MAC control signal eth_mode selects either MII or GMII. The MAC function asserts the eth_mode signal when the MAC operates in gigabit mode, which subsequently drives the MAC GMII to the PHY interface. The eth_mode signal is deasserted when the MAC function operates in 10/100 mode. In this mode, the MAC MII is driven to the PHY interface.

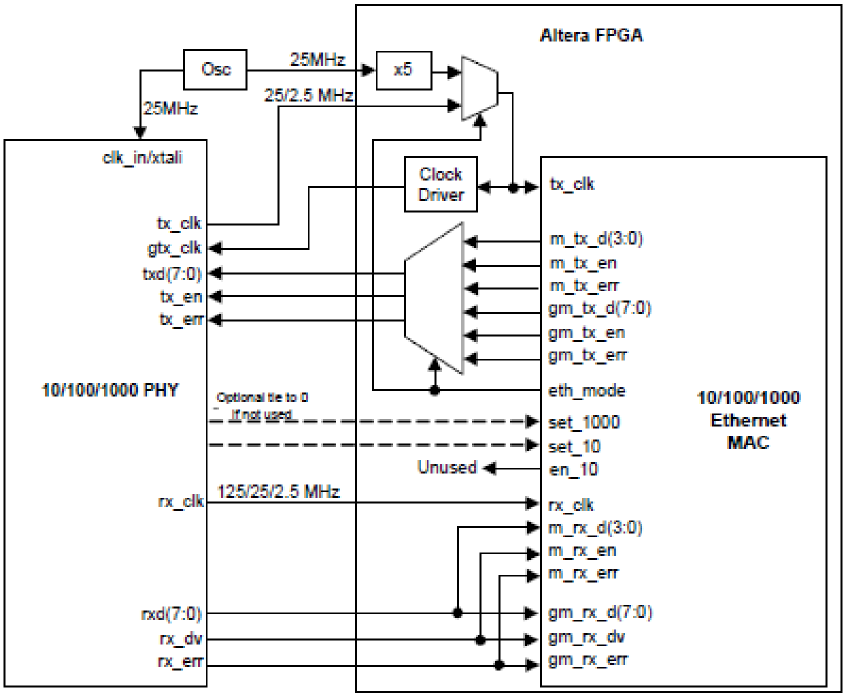


Figure 5. Programmable 10/100/1000Mbps Ethernet operation [8]

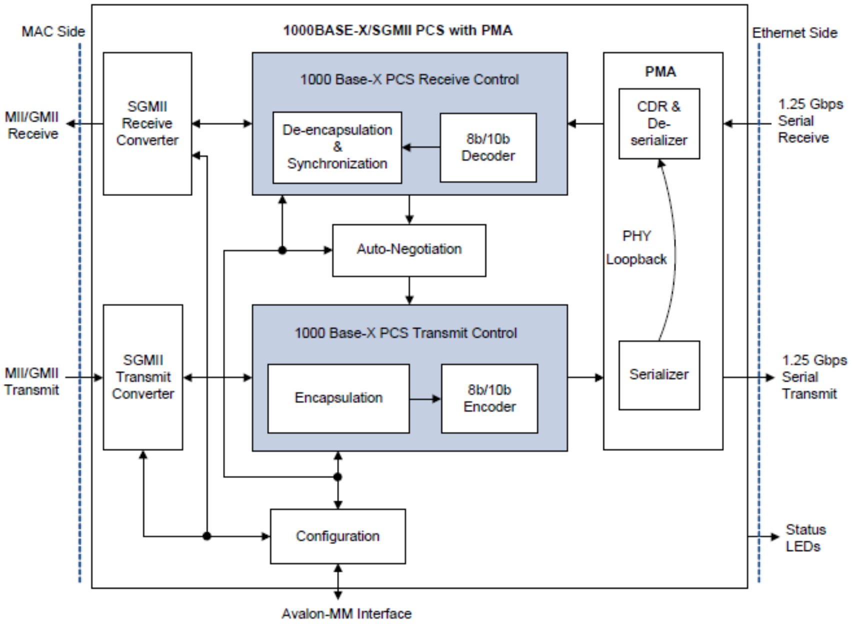


Figure 6 Block diagram of the PCS function with an embedded PMA [8]

The 1000BASE-X/SGMII (Serial GMII) PCS (Physical Coding Sublayer) function is accessible via GMII (1000BASE-X/SGMII) or MII (SGMII). Figure 6 gives the Block Diagram of the PCS function with an embedded PMA(Physical Medium Attachment).The PCS function interfaces to an on- or off-chip SERDES component via the industry standard Ten-Bit Interface (TBI). The PCS function can be configured with an embedded PMA. This configuration complies with the IEEE 802.3 Standard 1000BASE-X PMA specification [8]. PMA interoperates with an external Physical Medium Dependent (PMD) device, which drives the external copper or optical network. The interconnect between Altera and PMD devices can be TBI or 1.25 Gbps serial.

2.1 Transmit Operation

The transmit operation includes frame encapsulation and encoding.

2.1.1 Frame Encapsulation

The PCS function replaces the first preamble byte in the MAC frame with the start of frame /S/ symbol. Then, the PCS function encodes the rest of the bytes in the MAC frame with standard 8B/10B encoded characters. After the last FCS byte, the PCS function inserts the end of frame sequence, /T/ /R/ /R/ or /T/ /R/, depending on the number of character transmitted. Between frames, the PCS function transmits /I/ symbols. If the PCS function receives a frame from the MAC function with an error (gm_tx_err asserted during frame transmission), the PCS function encodes the error by inserting a /V/ character.

2.1.2 8b/10b Encoding

The 8B/10B encoder maps 8-bit words to 10-bit symbols to generate a DC balance and ensure disparity of the stream with a maximum run length of 5.

2.2 Receive Operation

The receive operation includes comma detection, decoding, de-encapsulation, synchronization, and carrier sense.

2.2.1 Comma Detection

The comma detection function searches for the 10-bit encoded comma character, K28.1/K28.5/K28.7, in consecutive samples received from PMA devices. When the K28.1/K28.5/K28.7 comma code group is detected, the PCS function realigns the data stream on a valid 10-bit character boundary. A standard 8b/10b decoder can subsequently decode the aligned stream. The comma detection function restarts the search for a valid comma character if the receive synchronization state machine loses the link synchronization.

2.2.2 8b/10b Decoding

The 8b/10b decoder performs the disparity checking to ensure DC balancing and produces a decoded 8-bit stream of data for the frame de-encapsulation function.

2.2.3 Frame De-encapsulation

The frame de-encapsulation state machine detects the start of frame when the /I/ /S/ sequence is received and replaces the /S/ with a preamble byte (0x55). It continues decoding the frame bytes and transmits them to the MAC function. The /T/ /R/ /R/ or the /T/ /R/ sequence is decoded as an end of frame. A /V/ character is decoded and sent to the MAC function as frame error. The state machine decodes sequences other than /I/ /I/ (Idle) or /I/ /S/ (Start of Frame) as wrong carrier.

During frame reception, the de-encapsulation state machine checks for invalid characters. When the state machine detects invalid characters, it indicates an error to the MAC function.

2.2.4 Synchronization

The link synchronization constantly monitors the decoded data stream and determines if the underlying receive channel is ready for operation. The link synchronization state machine acquires link synchronization if the state machine receives three code groups with comma consecutively without error. When link synchronization is acquired, the link synchronization state machine counts the number of invalid characters received. The state machine increments an internal error counter for each invalid character received and incorrectly positioned comma character. The internal error counter is decremented when four consecutive valid characters are received. When the counter reaches 4, the link synchronization is lost. The PCS function drives the led_link signal to 1 when link synchronization is acquired. This signal can be used as a common visual activity check using a board LED.

2.2.5 Carrier Sense

The carrier sense state machine detects an activity when the link synchronization is acquired and when the transmit and receive encapsulation or de-encapsulation state machines are not in the idle or error states. The carrier sense state machine drives the mii_rx_crs and led_crs signals to 1 when it detects an activity. The led_crs signal can be used as a common visual activity check using a board LED.

2.3 Collision Detection

A collision happens when non-idle frames are received from the PHY and transmitted to the PHY simultaneously. Collisions can be detected only in SGMII and half-duplex mode. When a collision happens, the collision detection state machine drives the mii_rx_col and led_col signals to 1. You can use the led_col signal as a visual check using a board LED.

2.4 SGMII Converter

You can enable the SGMII converter by setting the SGMII_ENA bit in the if_mode register to 1. When enabled and the USE_SGMII_AN bit in the if_mode register is set to 1, the SGMII converter is automatically configured with the capabilities advertised by the PHY. In 1000BASE-X mode, the PCS function always operates in gigabit mode and data duplication is disabled.

2.4.1 Transmit

In gigabit mode, the PCS and MAC functions must operate at the same rate. The transmit converter transmits each byte from the MAC function once to the PCS function. In 100-Mbps mode, the transmit converter replicates each byte received by the PCS function 10 times. In 10 Mbps, the transmit converter replicates each byte transmitted from the MAC function to the PCS function 100 times.

2.4.2 Receive

In gigabit mode, the PCS and MAC functions must operate at the same rate. The transmit converter transmits each byte from the PCS function once to the MAC function. In 100-Mbps mode, the receive converter transmits one byte out of 10 bytes received from the PCS function to the MAC. In 10-Mbps, the receive converter transmits one byte out of 100 bytes received from the PCS function to the MAC function.

2.5 Auto-Negotiation

Auto-negotiation is an optional function that can be started when link synchronization is acquired during system start up. To start auto-negotiation automatically, set the AUTO_NEGOTIATION_ENABLE bit in the PCS control register to 1. During auto-negotiation, the PCS function advertises its device features and exchanges them with a link partner device. If the SGMII_ENA bit in the if_mode register is set to 0, the PCS function operates in 1000BASE-X. Otherwise, the operating mode is SGMII.

2.6 Ten-bit Interface

In PCS variations embedded PMA, the PCS function implements a TBI to an external SERDES. On transmit, the SERDES must serialize tbi_tx_d[0], the least significant bit of the TBI output bus first and tbi_tx_d[9], the most significant bit of the TBI output bus last to ensure the remote node receives the data correctly On receive, the SERDES must serialize the TBI least significant bit first and the TBI most significant bit last

2.7 PHY Loopback

In PCS variations with embedded PMA targeting devices with GX transceivers, you can enable loopback on the serial interface to test the PCS and embedded PMA functions in isolation of the PMD. To enable loopback, set the sd_loopback bit in the PCS control register to 1.

2.8 PHY Power-Down

Power-down is controlled by the POWERDOWN bit in the PCS control register. When the PHY is in power-down state, the PCS function is in reset and any activities on the GMII transmit and the TBI receive interfaces are ignored. The management interface remains active and responds to management transactions from the MAC layer device.

2.9 Power-Down in PCS Variations with Embedded PMA

In PCS variations with embedded PMA targeting devices with GX transceivers, the power-down signal is internally connected to the power-down of the GX transceiver.

2.10 Reset

A hardware reset resets all logic synchronized to the respective clock domains whereas a software reset only resets the PCS state machines, comma detection function, and 8B10B encoder and decoder.To trigger a hardware reset, assert the respective reset signals: reg_clk (clk in PCS with embedded PMA core variations), tx_clk, and rx_clk. To trigger a software reset, set the RESET bit in the control register to 1.

3. FPGA IMPLEMENTATION

Figure 7 shows a high-level block diagram of the Triple Speed Ethernet (TSE) design[9]. The design includes two Altera TSE MegaCore functions (MAC + PCS + PMA) and is downloaded on Altera's Stratix II GX PCI Express Development Kit. There are two SFP (Small Form-factor Pluggable) cages built onto the kit. This design interfaces the TSE MegaCore function [8] with a Copper or Optical Fibre SFP module via a 1.25 Gbps serial transceiver that enables all 10, 100, and 1000 Mbps Ethernet operations. The design sends stream of Ethernet packets to the TSE MegaCore function, which can be looped back using SFP modules with an Ethernet fibre optic cable, copper cable or a switch.The design can demonstrate the operation of the TSE MegaCore function in various modes with live traffic upto the maximum throughput rate and show the error rate in the receiver. The design is built using Altera's Quartus II software and SOPC (System On Programmable Chip) builder.The Nios II processor is used as a control plane component for setting up and configuring the system components.

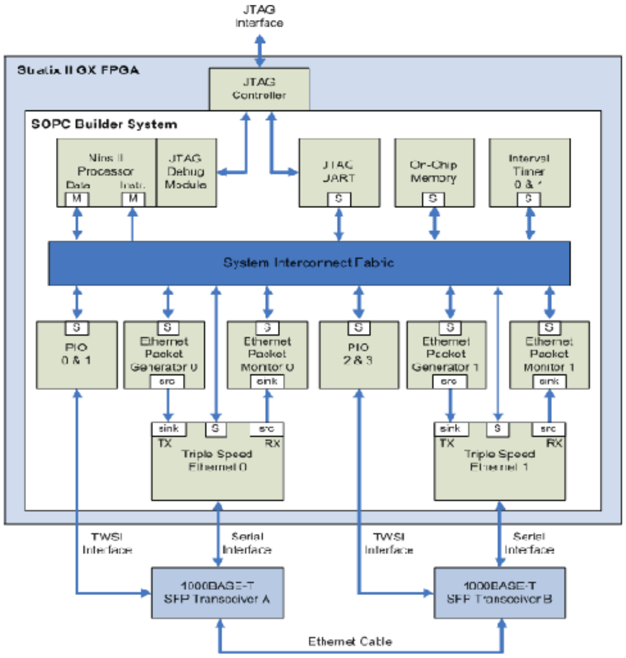


Figure 7.Block Diagram of Triple Speed Ethernet Reference Design [9]

Altera's Triple Speed Ethernet Design has been implemented and used as a platform for studying the performance of Gigabit Ethernet Standards 1000Base-LX, 1000Base-SX and 1000Base-T. The design has been implemented on Altera's Stratix II GX device EP2SGX90FF1508C3.Table 1 summarizes the resource utilization of this design.

Table . 1 Resource Utilisation of Triple Speed Ethernet Design

Parameter	Value	Utilization (%)
Logic	15281	21
Combinational ALUTs	11209	15
Dedicated logic registers	10419	14
Total pins	32	4
Total block memory bits	390494	9
DSP block 9-bit elements	8	2
Total PLLs	1	13
Total GXB Receiver Channels	2	13
Total GXB Transmitter Channels	2	13

4. Results and Discussion

The network performance instrument with measuring ability of full line rate is an important component of the system[10]. The Stratix® II GX PCI Express development board provides a hardware platform for developing and prototyping high-performance PCI Express (PCIe)-based designs[11].The Test System of the Ethernet design is given below in Figure 8. The Triple Speed Ethernet design is dumped onto the FPGA using Quartus II software and JTAG interface.

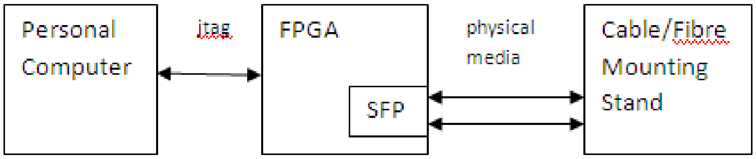


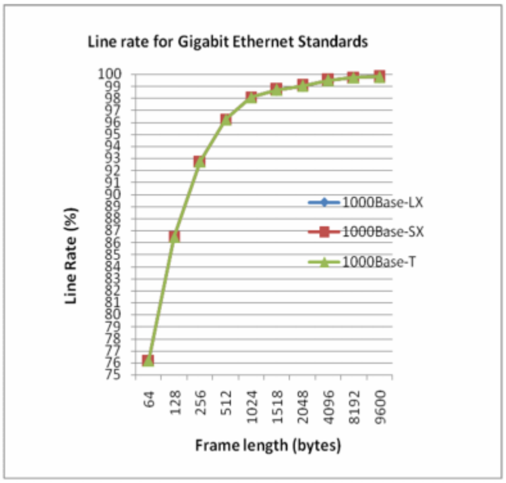
Figure 8. Test System for Performance Evaluation

The performance of above design was studied for Gigabit Ethernet standards 1000Base-LX ,1000Base-SX and 1000Base-T using various physical media such as Single Mode Fibre , Multimode Fibre and Copper cables (Cat5e) and corresponding SFP Transceivers. The system was tested by varying frame length and number of frames .The first test was performed by increasing the frame length and keeping the number of frames fixed . The test was repeated for two different values of number of frames i.e. 10^5 and 10^7 .The results of the tests performed for the three Gigabit Ethernet Standards 1000Base-LX ,1000Base-SX and 1000Base-T is given in Table 2. It is found that as the frame length is increased from 64 bytes to 9600 bytes, the line rate increases and achieves 99.79% for 9600 bytes. The throughput measured is 0 .76 Gbps for 64 bytes frame length and approaches 1 Gbps for 9600 bytes. The results are found to be nearly the same for the different Gigabit Ethernet standards.

Table 2. Line Rate and Throughput of the Network

Frame Length (bytes)	Line rate (%)	Throughput (bits/sec)
64	76.19	7.619E+08
128	86.48	8.649E+08
256	92.75	9.275E+08
512	96.24	9.624E+08
1024	98.08	9.808E+08
1518	98.70	9.870E+08
2048	99.03	9.903E+08
4096	99.51	9.951E+08
8192	99.75	9.975E+08
9600	99.79	9.979E+08

From the Figure 9 and Figure 10 ,it can be seen that the curves for the 3 standards 1000Base-LX ,1000Base-SX and 1000Base-T almost overlap. Hence the performance of the 3 standards is found to be almost similar.



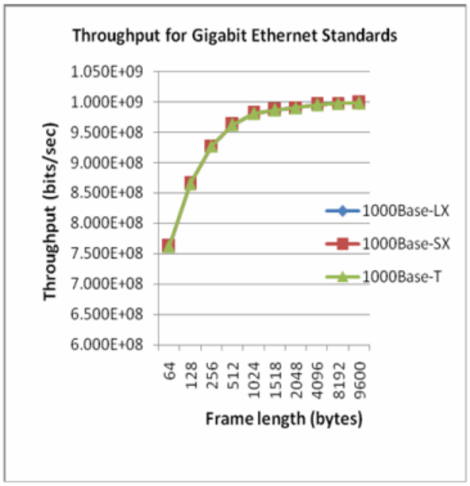


Figure 9. Line Rate vs Frame length

Figure 10. Throughput vs Frame length

The next test was performed by keeping the total number of bytes (N) sent constant. N =frame length×number of frames .The frame length was varied from 64 bytes to 9600 bytes and correspondingly number of frames was varied. The experiment was repeated for 3 different values of N i.e. 128×10^5, 128×10^6 and 128×10^7. Fig.11 shows that the throughput is the same for the 3 values of N for a particular frame length. Fig.12 shows that the line rate obtained is same for all the 3 values of N for a particular frame length. Also the values of Line rate and Throughput are the same as in Table 2 (where number of frames is kept fixed). Similarly ,Figure 13 shows that the Total transmission time(t) has reduced from .134s for 64 bytes frame length to .103s for

9600 bytes when N= 128×10^5,1.344s to 1.026s for N= 128×10^6 and 13.44 to 10.26s for N= 128×10^7 .The total transmission time remains almost constant from frame length 1024 bytes to 9600 bytes for all the three N values . This illustrates that there is an overhead for smaller size frames .Figure 14 shows that the total transmission time is directly proportional to the value of N . These tests were performed only for 1000Base-SX standard . The use of SOPC has given us the flexibility to choose from different software and hardware components and greatly reduce the system development cycle [12].

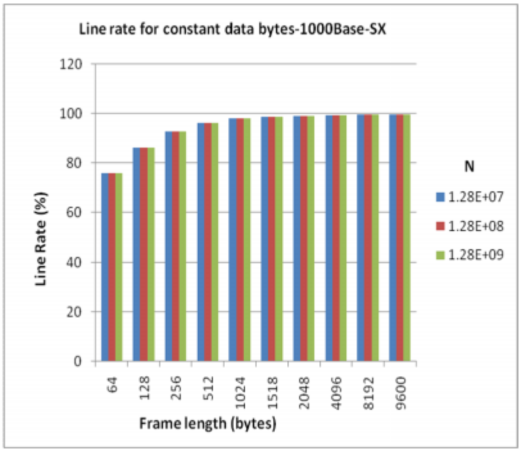


Figure 11. Line Rate vs Frame length

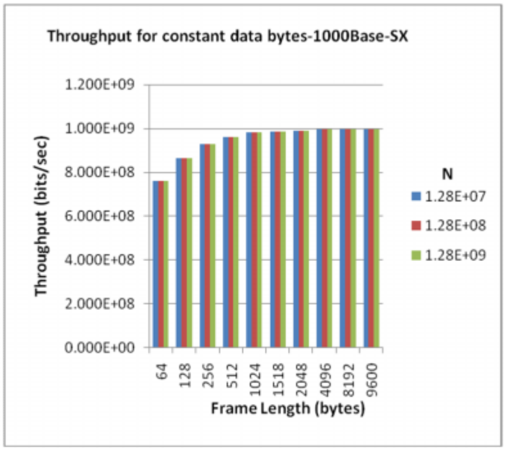


Figure 12. Throughput vs Frame length

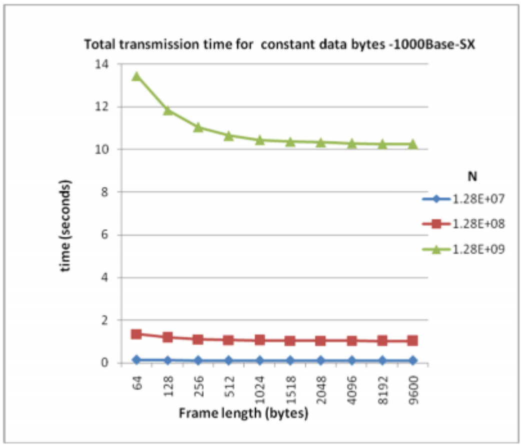


Figure 13. Transmission time vs Frame length

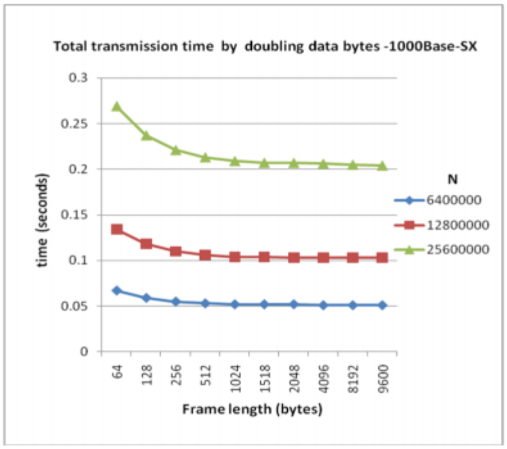


Figure 14. Transmission time vs Frame length

5. CONCLUSIONS

This paper gives an overview of the 10Mbps , 100Mbps and 1Gbps Ethernet technologies and describes its implementation on Altera's FPGA. Quartus II software is used to synthesize and create .sof file .The design is downloaded to the FPGA chip using JTAG interface.The resource utilization of this design is summarized.The performance of Gigabit Ethernet design is analysed and the line rate is found to be 76.19% for minimum 64 bytes packet size and approaches 100% for 9600 bytes frame length.The throughput is lowest for 64 bytes frame size and approaches

1Gbps for 9600 frame length .For a particular frame length, the throughput and line rate remain almost the same for all the different Gigabit Ethernet standards. Also, as the frame length increases , the total transmission time is found to be decreasing and remains constant from 1024 bytes frame length onwards.This design has been found to be very robust and we have also developed an experimental platform to introduce errors into the network and future work includes Error Detection and Correction Analysis using the same platform.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Altera Inc. USA for the MOU with Goa University and one of the authors V. R. Gad would like to thank University Grants Commission (UGC), New Delhi,India for providing FIP study leave under which this work is being carried out.

REFERENCES

[1] http://www.cisco.com

[2] Zou et al.,(2006) Journal of Zhejiang University SCIENCE A ISSN 1009-3095 (Print); ISSN 1862-1775 pp591-596

[3] Ciobotaru, M. et al,(2005)," Versatile FPGA-based Hardware Platform for Gigabit Ethernet Applications"

[4] Tinoosh Mohsenin , (2004) Rice University, "Design and Evaluation of FPGA-Based Gigabit-Ethernet/PCI Network Interface Card Thesis" .

[5] C. Kachris, et al., (2008) "Design and performance evaluation of an adaptive FPGA for network applications", Microelectron. J (2008) , doi: 10.1016/j.mejo. 2008.05.011 .

[6] Howard Frazier, (1998) "The 802.3z Gigabit Ethernet Standard" ,IEEE Network,May/June 1998,pp.6-7.

[7] IEEE Std 802.3™2008 , "Carrier sense multiple access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications"

[8] Triple-Speed Ethernet MegaCore Function User Guide © December 2010 Altera Corporation Application Note 483.

[9] Triple Speed Ethernet Data Path Reference Design AN-483-June 2009 ver. 1.1 Altera Corpora-tion.

[10] Meixia Duan, Hongling Han, (2011) "Research and Implementation of Gigabit Ethernet Full Line Rate", 2011 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference, pp. 736-738 .

[11] Stratix II GX PCI Express Development Board Reference Manual April 2007 Altera Corporation

[12] Yafang Wang, Cheng Zhang, Yanli Hou,Boning Hu, (2010) "Implementation of Gigabit Ethernet Network based on SOPC", Asia Pacific Conference on Wearable Computing Systems, pp.341-343.

Authors

V. R. Gad, M.Sc.**,** M.Phil.

Head , Dept. of Computer Science,G. V. M.'s G. G. P. R. College of Com. & Eco., Ponda, Goa, India. Graduated in Physics from Dhempe College, Miramar, Goa and completed M.Sc. (Electronics) from Goa University in 1994 and 1996 respectively and obtained M.Phil. in Electronics from Bharathidasan University, Tiruchirapalli in 2008.She has 12 years of teaching experience. She has worked on the University Grants Commission Minor Research Project "Design and Development of Computerised ID Card System" .She is a Research scholar in the Department of Electronics ,Goa University, Goa, India. Her current research interest includes Computer Networks , Error Control Coding and Embedded Systems.

Dr. R. S. Gad, M.Sc., Ph.D.

Reader, Dept. of Electronics, Goa University, Goa, India. Graduated in Physics from St. Xaviers College and completed M.Sc. Electronics, Department of Physics, Goa University, Goa, India.He has worked on the Indian Council of Medical Research and University Grant Commission, New Delhi, funded research project in the area of non-invasive glucometer. Also closely associated with the Million Book Project of Carnegie Mellon University, USA and related digital repository projects of Indian Navy. Associated with ALTERA Inc. USA under the MOU with ALTERA University program. Attended summer training at CEDT IISc, Bangalore for two months from April 27, 1998. Attended SERC School on Bio-ph otonics, CAT Indore: February 06, 2006 -February 24, 2006; supported by DST, New Delhi. Dr. Gad , is Sponsored and Administered "Leading Engineer of the World 2008" & 2000 Outstanding Intellectuals of the 21st century 2009/2010, by International Biographical Center, Cambridge, England. Dr. Gad, was a winner in Mentor Graphics Design contest ' Design and verification of LC3 processor' for year 2010 in India. He is also recipient of the Indian National Science Academy Fellowship for the year 2012-13.

Prof. G. M. Naik , M.Sc., Ph.D.

Professor & Head, Dept. of Electronics, Goa University, Goa, India Graduated from Karnataka University in Physics and Master in Electronics from Gulburga University. Prof. Gourish Naik obtained his Ph.D (Physics) from Indian Institute of Science, Bangalore (1987) and served the institute as research associate in the areas of Optoelectronics and Communication till 1993. For the last 15 years, he is associated with Goa University Electronics Program. He is the founder Head of University Instrumentation Center He is also coordinator of DEITI (an educational broadcast studio supported by Indian Space Research). Has co- authored two books on Embedded Systems and Programming published by Springer (Holland). Presently he is head of Dept. of Electronics at Goa University.

Books: 1)J. S. Parab, G. M. Naik et al ; "Exploring C For Microcontrollers: A Hands On Approach" Publisher: Springer Verlag, 2008.2) J. S. Parab, G. M. Naik et al ; "Practical Aspects Of Embedded System Design Using Microcontrollers" Publisher: Springer Verlag, 2008.