

TOE100G-IP 4-Session with CPU reference design

Rev1.1 21-Mar-22

1 Introduction

The default TOE100G-IP reference design implements one TOE100G-IP for transferring TCP payload data by using one session. It is found that the performance is limited when the test environment is FPGA and PC, not FPGA and FPGA. The maximum performance when transferring with the test application run on PC is about 3000 – 4000 Mbyte/s.

When the data is transferred by using multiple sessions shared the same 100G Ethernet channel, the total performance between FPGA and PC is much increased. For example, when using four sessions by four TOE100G-IPs, the total performance of four sessions is about 8000 – 9000 Mbyte/s.

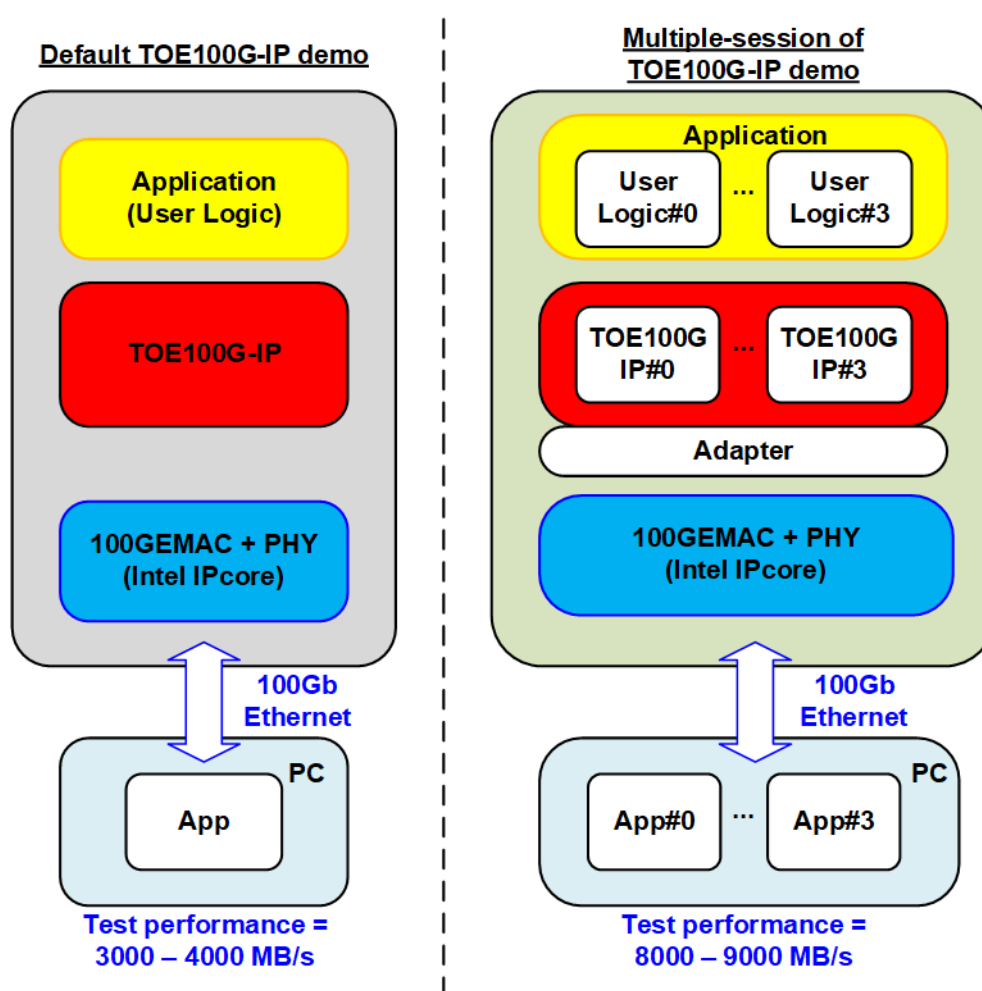


Figure 1-1 Multiple-session comparison with Uni-session

This document describes the reference design that includes four TOE100G-IPs that share the same 100GEMAC and PHY to support four sessions. The adapter is designed to interface four TOE100G-IPs and one 100GEMAC. User logic and Test application for running multiple-session demo are almost similar to one-session demo, except the test application of full-duplex demo. “tcp_client_txrx_single” is applied to run on PC instead. This application is run the test for one round, not forever loop like “tcp_client_txrx_40G”.

Though the multiple-session demo implements four sessions, the user can enable each TOE100G-IP independently to check the performance or the operation when using less than four sessions. Also, the transfer direction of each session can be configured independently. The user can modify the multiple-session reference design to increase the number of sessions to match the system requirement.

When the test environment is FPGA and FPGA, the performance of one session and four sessions is not much different. According to the test result of the default TOE100G-IP demo, using one session by FPGA and FPGA achieves the maximum performance of 100Gb Ethernet (about 12000 Mbyte/s). Therefore, using multiple sessions cannot improve the performance and may slightly reduce the performance from overhead time to switch the session. However, using multiple sessions may be applied when the system needs to transfer the data from many sources which are run independently.

2 Hardware overview

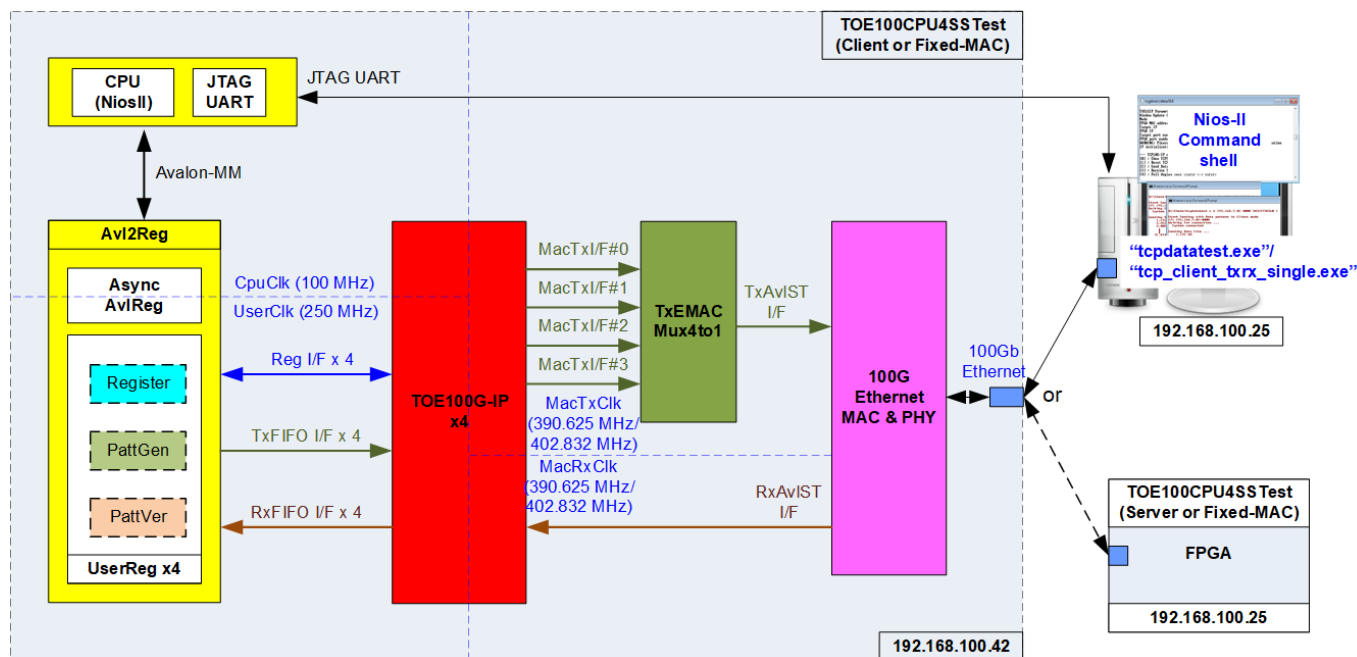


Figure 2-1 Demo block diagram

In test environment, two devices are used for 100Gb Ethernet transferring. When using FPGA and PC, FPGA is initialized by client mode and PC is initialized by server mode. On the other hand, when using two FPGAs, it may be initialized by Client \leftrightarrow Server, Client \leftrightarrow Fixed-MAC, or Fixed-MAC \leftrightarrow Fixed-MAC. Two test applications are called on PC for transferring the data, tcpdatatest or tcp_client_txrx_single.

In FPGA system, four TOE100G-IPs are included. Four MacTxI/F which are output from TOE100G-IPs are connected to 100G EMAC and PHY via TxEMACMux4to1 module. While Rx interface of 100G EMAC and PHY is fed to four TOE100G-IPs directly. User interface of four TOE100G-IP are mapped to Avl2Reg which consists of AsyncAvlReg and four UserRegs, one UserReg connecting to one TOE100G-IP independently. There are three logic groups inside UserReg, i.e., Register file for interfacing with Register interface, PattGen for sending test data via Tx FIFO interface, and PattVer for verifying test data via Rx FIFO interface. Register files of UserReg are controlled by CPU firmware through Avalon-MM bus.

There are four clock domains in the design, i.e., CpuClk which is the clock for running the CPU system, MacTxClk and MacRxClk which are the clock output from 100G EMAC for synchronous with Tx and Rx interface of EMAC IP respectively, and UserClk which is the clock for running user logic of TOE100G-IP. In real system, the user can change the frequency of CpuClk and UserClk. According to TOE100G-IP datasheet, clock frequency of UserClk must be more than or equal to 220 MHz.

Note: When using 100G EMAC Soft IP, MacTxClk and MacRxClk have the same frequency, 390.625 MHz, but it is different clock. While using 100G EMAC Hard IP, the same clock at 402.832 MHz is applied for both Tx and Rx interface.

AsyncAvlReg is designed to support asynchronous signals between CpuClk and UserClk. More details of each module are described as follows.

2.1 100G Ethernet MAC and PHY

For Ethernet IP core (Soft IP), the user interface to connect TOE100G-IP is 512-bit Avalon stream bus which runs at 390.625 MHz. Though Tx and Rx interface of Avalon stream bus are run at 390.625 MHz, the clock domain of Tx and Rx interface are different. More details of 100G Ethernet MAC IP are described in following link.

<https://www.intel.com/content/www/us/en/programmable/products/intellectual-property/ip/interface-protocols/m-alt-ll100gb-ethernet.html>

For Hard Ethernet IP core, the user interface to connect TOE100G-IP is 512-bit Avalon stream bus which runs at 402.832 MHz. The Tx and Rx interface are synchronous to the same clock domain. More details of 100G Ethernet MAC IP are described in following link

<https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/ug/ug-s10-etile-hip-ethernet.pdf>

In the reference design, the physical interface of 100G Ethernet uses 100GBASE-R standard.

2.2 TOE100G-IP

TOE100G-IP implements TCP/IP stack and offload engine. User interface has two signal groups, i.e., control signals and data signals. Register interface is applied to set control registers and monitor status signals. Data signals are accessed by using FIFO interface. The interface with 100G EMAC is 512-bit Avalon-ST interface.

More details are described in datasheet.

https://dgway.com/products/IP/TOE100G-IP/dg_toe100gip_data_sheet_intel.pdf

2.3 TxMuxEMAC4to1

The system consists of four TOE100G-IPs which share the same EMAC. Therefore, TxEMACMux4to1 is designed to transfer the transmitted packet from four TOE100G-IPs to one EMAC. The core signal inside this module is rChSel that is applied to select one active TOE100G-IP module from four modules. Timing diagram of this module is displayed in Figure 2-2.

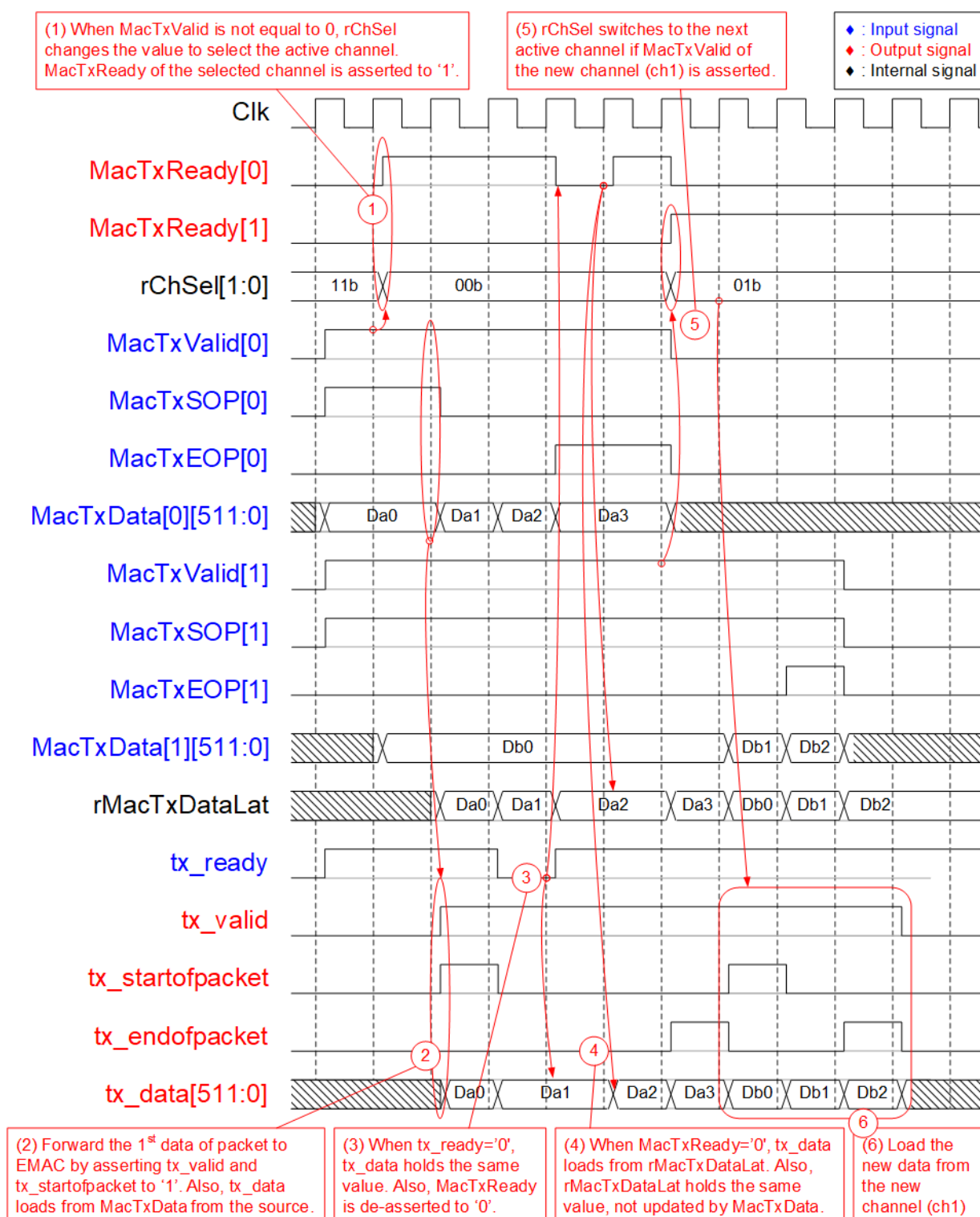


Figure 2-2 TxEMACMux4to1 timing diagram

- 1) If current channel (Ch#3) does not transfer the data and the new channel requests to transfer data by asserting MacTxValid to '1', rChSel (the signal to indicate the active channel) will switch the value to the new channel. In Figure 2-2, there are two channels that send the request at the same time. The selected channel will be the nearest channel of the order 0 -> 1 -> 2 -> 3 -> 0. Therefore, Ch#0 is higher priority than Ch#1 when the current channel is 3. MacTxReady of the selected channel (Ch#0) is asserted to '1' to accept the first data.
- 2) The input signals of the selected channel (Ch#0), i.e., MacTxSOP (start-of-packet), MacTxEOP (end-of-packet), and MacTxData (512-bit data) are loaded to the output signals of EMAC (tx_startofpacket, tx_endofpacket, and tx_data respectively). Also, tx_valid is asserted to '1' to start sending the new packet to EMAC.
- 3) When EMAC is not ready to receive data by de-asserting tx_ready to '0', all output signals of EMAC hold the same value. Also, MacTxReady of the active channel is de-asserted to '0' to hold the input signals from the source.
- 4) After EMAC re-asserts tx_ready to accept the data, the next output signals to EMAC will be loaded from the internal latch register (rMacTxDataLat). The internal latch register is designed to load the data from the active source when MacTxReady is asserted to '1'. Therefore, the latch register is the temporal buffer for storing the unsent data to EMAC when EMAC pauses data transmission.
- 5) After accepting the final data of a packet from the active channel, the next active channel is scanned. If MacTxValid of remaining channels is asserted, rChSel will update the value following the rule (0 -> 1 -> 2 -> 3 -> 0). In Figure 2-2, the next active channel is Ch#1, so rChSel is set to 01b to accept the data from Ch#1.
- 6) The input signals (MacTxSOP, MacTxEOP, and MacTxData) of the active channel (Ch#1) are forwarded to be the output signals of EMAC (tx_startofpacket, tx_endofpacket, and tx_data) until transferring the final data of a packet.

2.4 CPU and Peripherals

32-bit Avalon-MM is applied to be the bus interface for the CPU accessing the peripherals such as Timer and JTAG UART. To control and monitor the test system, the control and status signals are connected to register for CPU access as a peripheral through 32-bit Avalon-MM bus. CPU assigns the different base address and the address range to each peripheral for accessing one peripheral at a time.

In the reference design, the CPU system is built with one additional peripheral to access the test logic. So, the hardware logic must be designed to support Avalon-MM bus standard for supporting CPU writing and reading. Avl2Reg module is designed to connect the CPU system as shown in Figure 2-3.

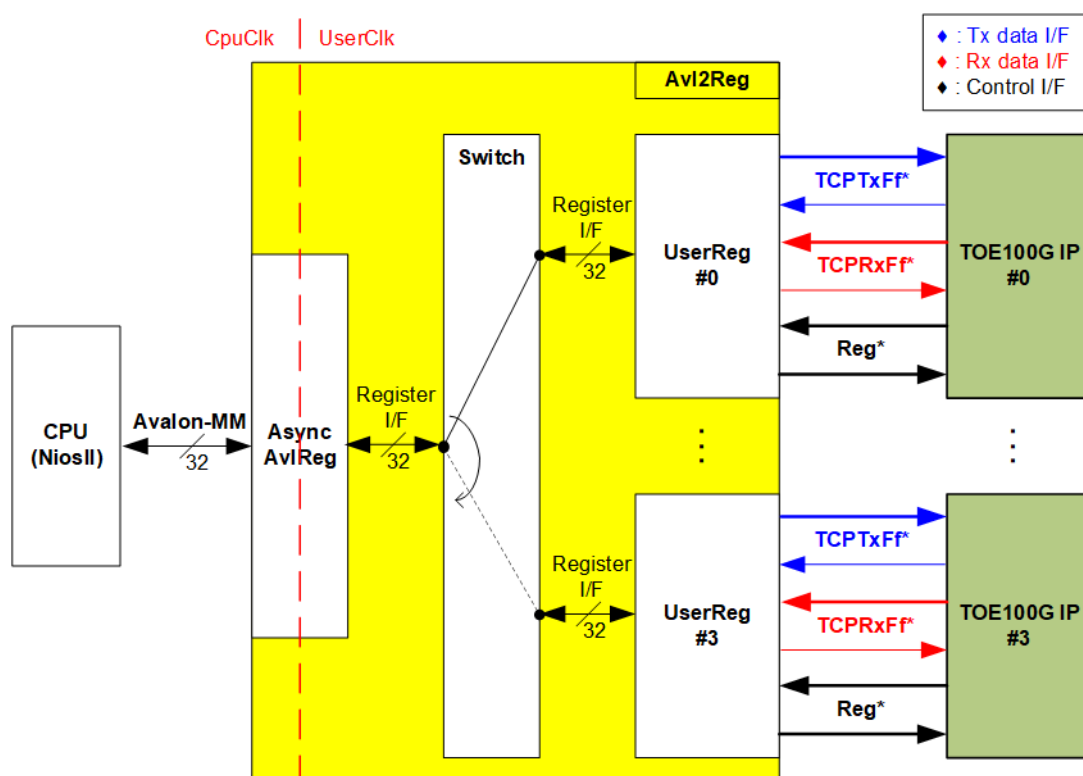


Figure 2-3 Avl2Reg block diagram

Similar to the default TOE100G-IP reference design, Avl2Reg consists of AsyncAvlReg and UserReg. To support multiple sessions, four UserRegs are integrated. Therefore, switch logic is designed to decode the address requested by AsyncAvl2Reg. Two upper bits are applied for selecting the active channel. When two upper bits of the requested address are equal to "00", "01", "10", or "11", switch logic connects Register I/F to UserReg#0, #1, #2, or #3 respectively. Switch logic includes D Flip-Flop for processing, so latency time for write and read access is increased, comparing to the default TOE100G-IP.

AsyncAvlReg and UserReg are the same module applied in the default TOE100G-IP reference design. AsyncAvlReg converts the Avalon-MM signals to be the simple register and includes asynchronous logic to support clock domain crossing between CpuClk domain and UserClk domain. UserReg is the example of user logic to interface with TOE100G-IP. More details of AsyncAvlReg and UserReg are described as follows.

2.4.1 AsyncAvlReg

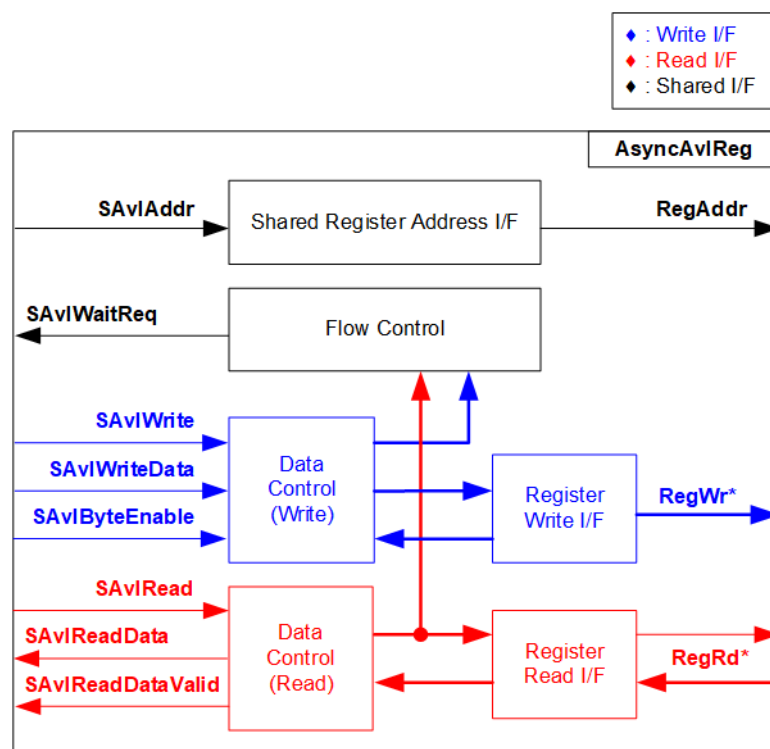


Figure 2-4 AsyncAvlReg interface

The signal on Avalon-MM bus interface can be split into three groups, i.e., Write channel (blue color), Read channel (red color), and Shared control channel (black color). More details of Avalon-MM interface specification are described in following document.

https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/manual/mnl_avalon_spec.pdf

According to Avalon-MM specification, one command (write or read) can be operated at a time. The logics inside AsyncAvlReg are split into three groups, i.e., Write control logic, Read control logic and Flow control logic. Flow control logic asserts SAvlWaitReq to hold the next request from Avalon-MM interface if the current request does not finish. Write control and Write data I/F of Avalon-MM bus are latched and transferred to be Write register interface with clock domain crossing registers. Similarly, Read control I/F are latched and transferred to be Read register interface. After that, the returned data from Register Read I/F is transferred to Avalon-MM bus by using clock domain crossing registers. Address I/F of Avalon-MM is latched and transferred to Address register interface as well.

The simple register interface is compatible with single-port RAM interface for write transaction. The read transaction of the register interface is slightly modified from RAM interface by adding RdReq and RdValid signals for controlling read latency time. The address of register interface is shared for write and read transaction, so user cannot write and read the register at the same time. The timing diagram of the register interface is shown in Figure 2-5

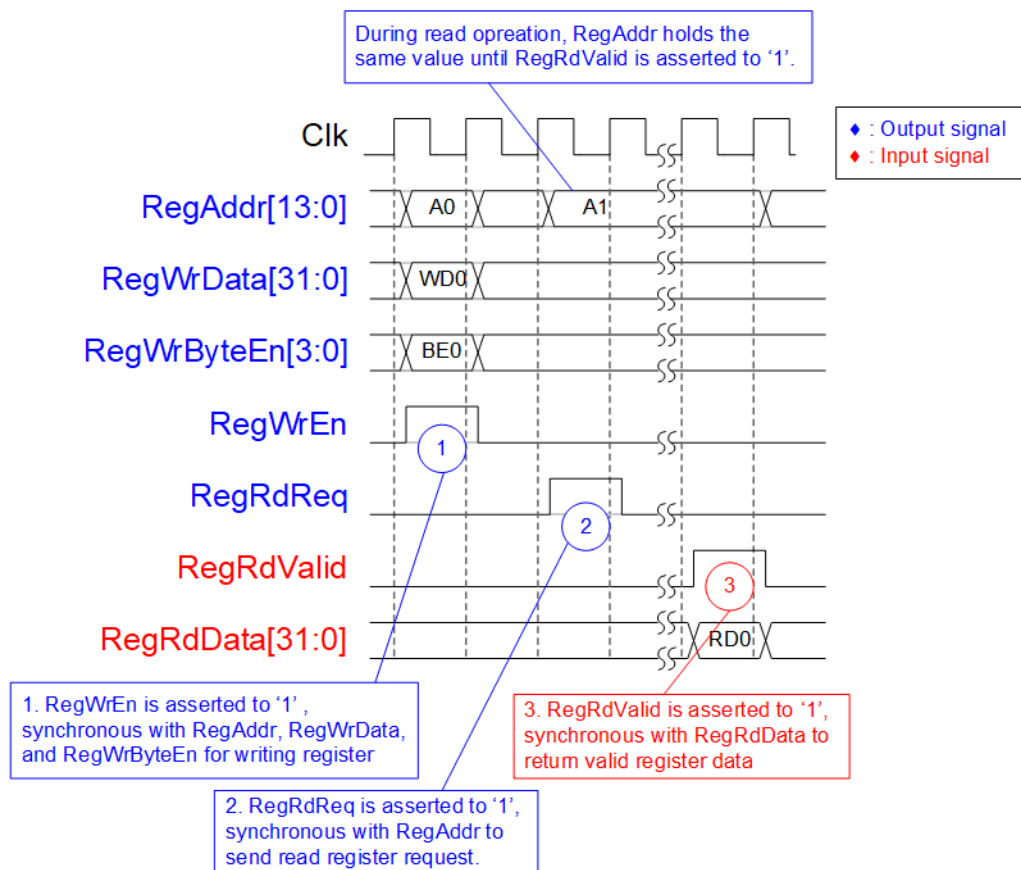


Figure 2-5 Register interface timing diagram

- 1) To write register, the timing diagram is similar to single-port RAM interface. RegWrEn is asserted to '1' with the valid signal of RegAddr (Register address in 32-bit unit), RegWrData (write data of the register), and RegWrByteEn (the write byte enable). Byte enable has four bits to be the byte data valid. Bit[0], [1], [2], and [3] are equal to '1' when RegWrData[7:0], [15:8], [23:16], and [31:24] are valid respectively.
- 2) To read register, AsyncAvlReg asserts RegRdReq to '1' with the valid value of RegAddr. 32-bit data must be returned after receiving the read request. The slave must monitor RegRdReq signal to start the read transaction. During read operation, the address value (RegAddr) does not change the value until RegRdValid is asserted to '1'. So, the address can be used for selecting the returned data by using multiple layers of multiplexer.
- 3) The read data is returned on RegRdData bus by the slave with asserting RegRdValid to '1'. After that, AsyncAvlReg forwards the read data to SAVlRead interface.

2.4.2 UserReg

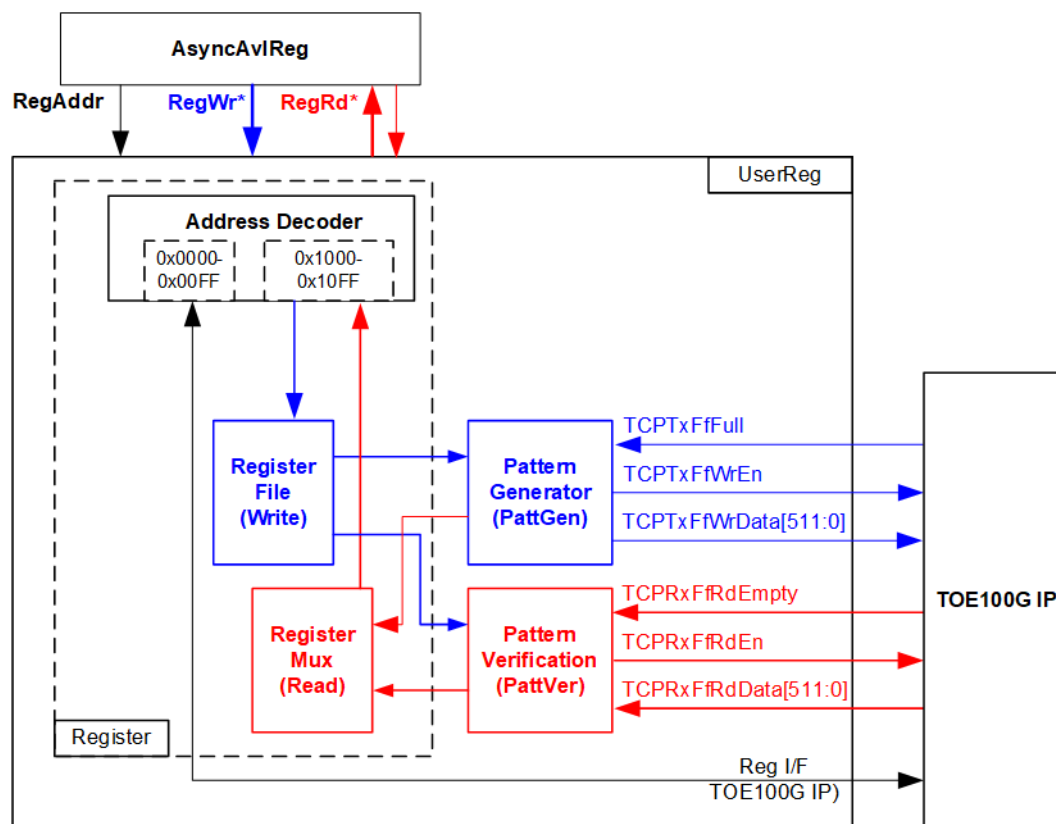


Figure 2-6 UserReg block diagram

The logic inside UserReg has three operations, i.e., Register, Pattern generator (PattGen), and Pattern verification (PattVer). Register block decodes the address requested from AsyncAvlReg and then selects the active register for write or read transaction. Pattern generator block is designed to send 512-bit test data to TOE100G-IP following FIFO interface standard. Pattern verification block is designed to read and verify 512-bit data from TOE100G-IP following FIFO interface standard. More details of each block are described as follows.

Register Block

Each UserReg uses the address area from 0x0000 – 0x1FFF. To support four sessions, four UserReg modules are mapped to 0x0000 – 0x7FFF by following assignment.

- 0x0000 – 0x1FFF: TOE100G-IP#0 and UserReg#0
- 0x2000 – 0x3FFF: TOE100G-IP#1 and UserReg#1
- 0x4000 – 0x5FFF: TOE100G-IP#2 and UserReg#2
- 0x6000 – 0x7FFF: TOE100G-IP#3 and UserReg#3

The address is split into two areas in each session, i.e., TOE100G-IP register (0x0000-0x00FF) and UserReg register (0x1000-0x10FF). Therefore, the upper bit of RegAddr is applied to select the active area between TOE100G-IP or UserReg register. While the lower bits of RegAddr are fed to TOE100G-IP and internal registers of UserReg to select the active register. The registers inside UserReg are 32-bit data, so write byte enable (RegWrByteEn) is not used. To write hardware registers, the CPU must use 32-bit pointer to place 32-bit valid value on the write data bus.

To read register, one multiplexer is included within UserReg to select the read data. Totally, the latency time to read data is equal to one clock cycle, so RegRdValid is created by RegRdReq with asserting one D Flip-flop. However, there is switch logic inside Avl2Reg which increases latency time of read access about two clock cycles. More details of the address mapping within UserReg module are shown in Table 2-1

Table 2-1 Register map Definition

Address	Register Name	Description
Wr/Rd	(Label in the "toe100g4ssstest.c")	
BA+0x0000 – BA+0x00FF: TOE100G-IP#0 Register Area		
More details of each register are described in TOE100G-IP datasheet.		
BA+0x0000	TOE_RST_REG	Mapped to RST register within TOE100G-IP#0
BA+0x0004	TOE_CMD_REG	Mapped to CMD register within TOE100G-IP#0
BA+0x0008	TOE_SML_REG	Mapped to SML register within TOE100G-IP#0
BA+0x000C	TOE_SMH_REG	Mapped to SMH register within TOE100G-IP#0
BA+0x0010	TOE_DIP_REG	Mapped to DIP register within TOE100G-IP#0
BA+0x0014	TOE_SIP_REG	Mapped to SIP register within TOE100G-IP#0
BA+0x0018	TOE_DPN_REG	Mapped to DPN register within TOE100G-IP#0
BA+0x001C	TOE_SPN_REG	Mapped to SPN register within TOE100G-IP#0
BA+0x0020	TOE_TDL_REG	Mapped to TDL register within TOE100G-IP#0
BA+0x0024	TOE_TMO_REG	Mapped to TMO register within TOE100G-IP#0
BA+0x0028	TOE_PKL_REG	Mapped to PKL register within TOE100G-IP#0
BA+0x002C	TOE_PSH_REG	Mapped to PSH register within TOE100G-IP#0
BA+0x0030	TOE_WIN_REG	Mapped to WIN register within TOE100G-IP#0
BA+0x0034	TOE_ETL_REG	Mapped to ETL register within TOE100G-IP#0
BA+0x0038	TOE_SRV_REG	Mapped to SRV register within TOE100G-IP#0
BA+0x003C	TOE_VER_REG	Mapped to VER register within TOE100G-IP#0
BA+0x0040	TOE_DML_REG	Mapped to DML register within TOE100G-IP#0
BA+0x0044	TOE_DMH_REG	Mapped to DMH register within TOE100G-IP#0
BA+0x1000 – BA+0x10FF: UserReg#0 Control/Status		
BA+0x1000	Total transmit length	Wr [31:0] – Total amount of transmitted data of UserReg#0 in 512-bit unit.
Wr/Rd	USER_TXLEN_REG	Valid from 1-0xFFFFFFFF. Rd [31:0] – Current amount of transmitted data of UserReg#0 in 512-bit unit. The value is cleared to 0 when USER_CMD_REG is written by user.
BA+0x1004	User Command	Wr
Wr/Rd	USER_CMD_REG	[0] – Set '0' to start transmitting data from UserReg#0. [1] – Data verification enable ('0': Disable data verification, '1': Enable data verification) Rd [0] – Busy of PattGen inside UserReg#0 ('0': Idle, '1': PattGen is busy) [1] – Data verification error ('0': Normal, '1': Error) This bit is auto-cleared when user starts new operation or reset. [2] – Mapped to ConnOn signal of TOE100G-IP#0
BA+0x1008	User Reset	Wr
Wr/Rd	USER_RST_REG	[0] – Reset signal. Set '1' to reset UserReg#0. This bit is auto-cleared to '0'. [8] – Set '1' to clear TimerInt latched value of TOE100G-IP#0 Rd [8] – Latched value of TimerInt output from TOE100G-IP#0 ('0': Normal, '1': TimerInt is asserted) This flag is cleared by system reset condition or setting USER_RST_REG[8]='1'. [16] – Ethernet linkup status from Ethernet MAC ('0': Not linkup, '1': Linkup)

Address	Register Name	Description
Wr/Rd	(Label in the "toe100g4ssstest.c")	
BA+0x1000 – BA+0x10FF: UserReg#0 Control/Status		
BA+0x100C Rd	FIFO status USER_FFSTS_REG	Rd[5:0] - Mapped to TCPRxFfLastRdCnt signal of TOE100G-IP#0 [15:6] - Mapped to TCPRxFfRdCnt signal of TOE100G-IP#0 [24] - Mapped to TCPTxFfFull signal of TOE100G-IP#0
BA+0x1010 Rd	Total receive length USER_RXLEN_REG	Rd[31:0] – Current amount of received data from TOE100G-IP#0 in 512-bit unit. The value is cleared to 0 when USER_CMD_REG is written by user.
BA+0x1080 Rd	EMAC IP version EMAC_VER_REG	Rd[31:0] – Mapped to IPVersion output from DG EMAC-IP when the system integrates DG EMAC-IP. In this demo, it is equal to 0.
BA+0x2000 – BA+0x20FF: TOE100G-IP#1 Register Area BA+0x3000 – BA+0x30FF: UserReg#1 Control/Status BA+0x4000 – BA+0x40FF: TOE100G-IP#2 Register Area BA+0x5000 – BA+0x50FF: UserReg#2 Control/Status BA+0x6000 – BA+0x60FF: TOE100G-IP#3 Register Area BA+0x7000 – BA+0x70FF: UserReg#3 Control/Status		
BA+0x2000- BA+0x20FF	TOE100G-IP#1 register area, similar to BA+0x0000 – BA+0x00FF	
BA+0x3000- BA+0x30FF	UserReg#1 control/status, similar to BA+0x1000 – BA+0x10FF	
BA+0x4000- BA+0x40FF	TOE100G-IP#2 register area, similar to BA+0x0000 – BA+0x00FF	
BA+0x5000- BA+0x50FF	UserReg#2 control/status, similar to BA+0x1000 – BA+0x10FF	
BA+0x6000- BA+0x60FF	TOE100G-IP#3 register area, similar to BA+0x0000 – BA+0x00FF	
BA+0x7000- BA+0x70FF	UserReg#3 control/status, similar to BA+0x1000 – BA+0x10FF	

Pattern Generator

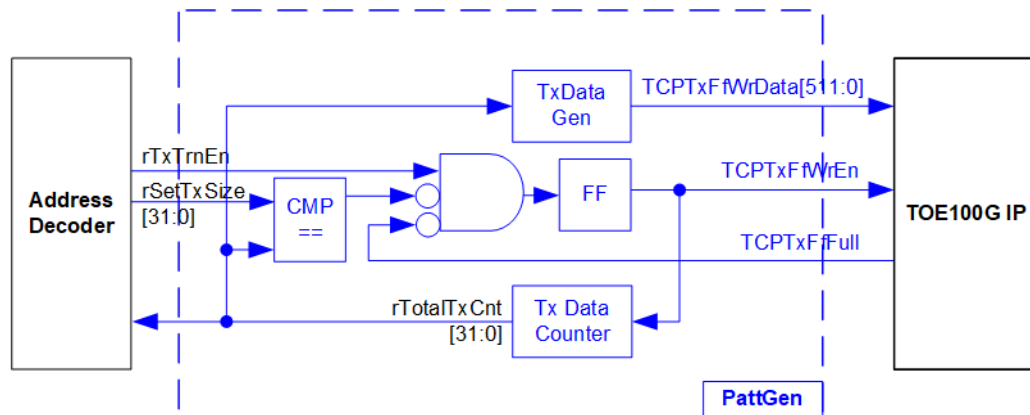


Figure 2-7 PattGen block

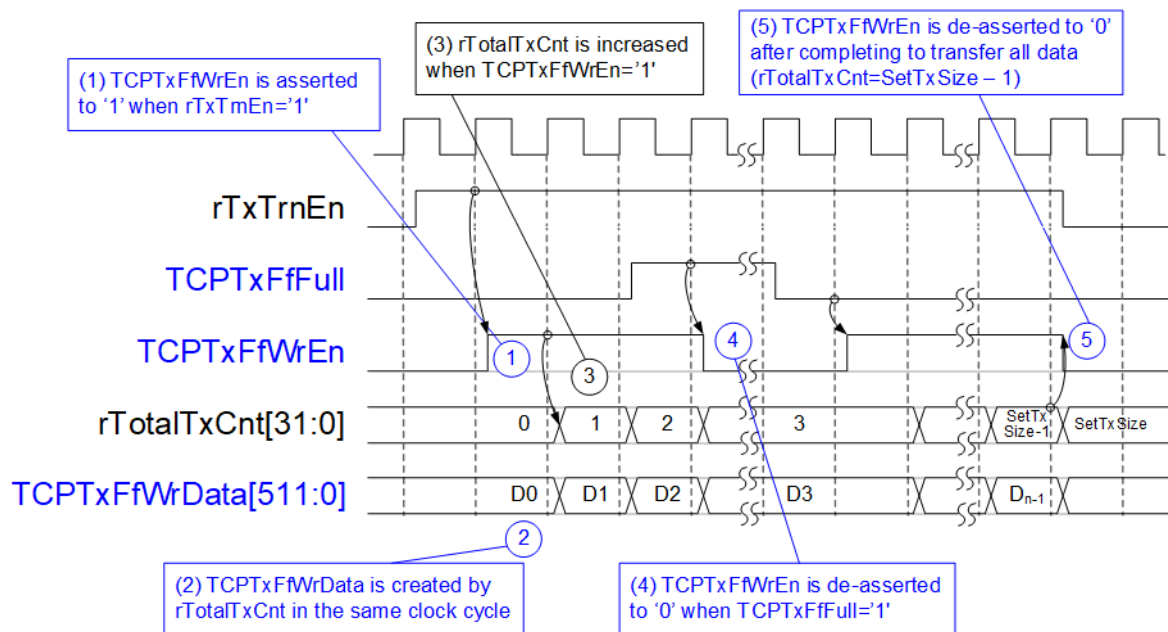


Figure 2-8 PattGen timing diagram

Figure 2-7 shows the details of PattGen which generates test data to TOE100G-IP. Timing diagram to show the relation of each logic is displayed in Figure 2-8.

To start PattGen operation, the user sets `USER_CMD_REG[0]='0'` and then `rTxTrnEn` is asserted to '1'. When `rTxTrnEn` is '1', `TCPTxFfWrEn` is controlled by `TCPTxFfFull`. `TCPTxFfWrEn` is de-asserted to '0' when `TCPTxFfFull` is '1'. `rTotalTxCnt` is the data counter to check total amount of transmitted data to TOE100G-IP. Also, `rTotalTxCnt` is used to generate 32-bit incremental data for `TCPTxFfWrData` signal. After all data is transferred completely (Total amount of data is equal to `rSetTxSize-1`), `rTxTrnEn` is de-asserted to '0'.

Pattern Verification

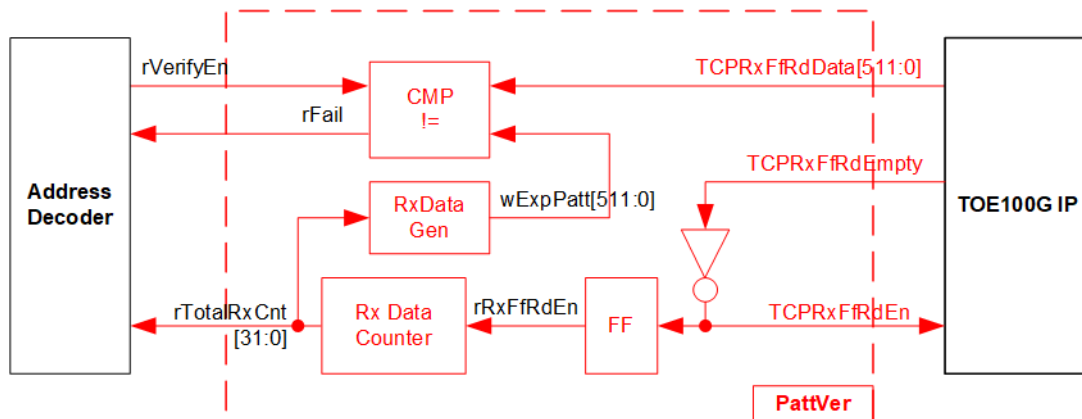


Figure 2-9 PattVer block

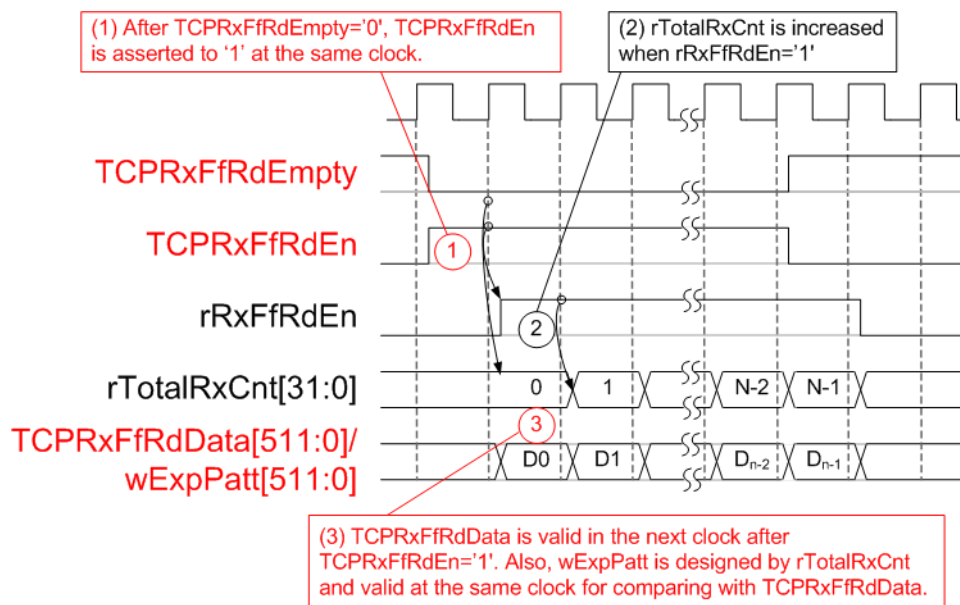


Figure 2-10 PattVer Timing diagram

Figure 2-9 shows the details of PattVer logic for reading the data from TOE100G-IP with or without data verification, controlled by rVerifyEn flag which is set by the user. Timing diagram of the logic is displayed in Figure 2-10.

When rVerifyEn is set to '1', data comparison is enabled to compare read data (TCPRxFfRdData) with the expected pattern (wExpPatt). If data verification is failed, rFail is asserted to '1'. TCPRxFfRdEn is designed by using NOT logic of TCPRxFfRdEmpty. TCPRxFfRdData is valid for data comparison in the next clock. rRxFfRdEn, one clock latency of TCPRxFfRdEn, is applied to be counter enable of rTotalRxCnt, counting total amount of received data. rTotalRxCnt is used to generate wExpPatt, so wExpPatt is valid at the same time as TCPRxFfRdData valid.

3 CPU Firmware on FPGA

For running the multisession reference design and further test modification, user should realize the information below.

- 1) Two end points of the communication are referred, i.e., FPGA#0 (left) and the target (right). The FPGA#0 always contains four TOE100G-IPs while the target can be PC with test application or the target can be FPGA#1 with TOE100G-IP(s) as shown in Figure 2-1.
- 2) In FPGA#0, up to four TCP ports/sessions are supported by using four TOE100G-IPs with the same FPGA#0 IP address. It does not need to run all four TOE100G-IPs. For example, one TOE100G-IP is not used when running three sessions in FPGA#0.
- 3) In the target, test system can be applied by using multiple targets to communicate with FPGA#0 when Ethernet switch is added to the system. Each target has unique IP address. However, the total amount of session from all target is up to four sessions.

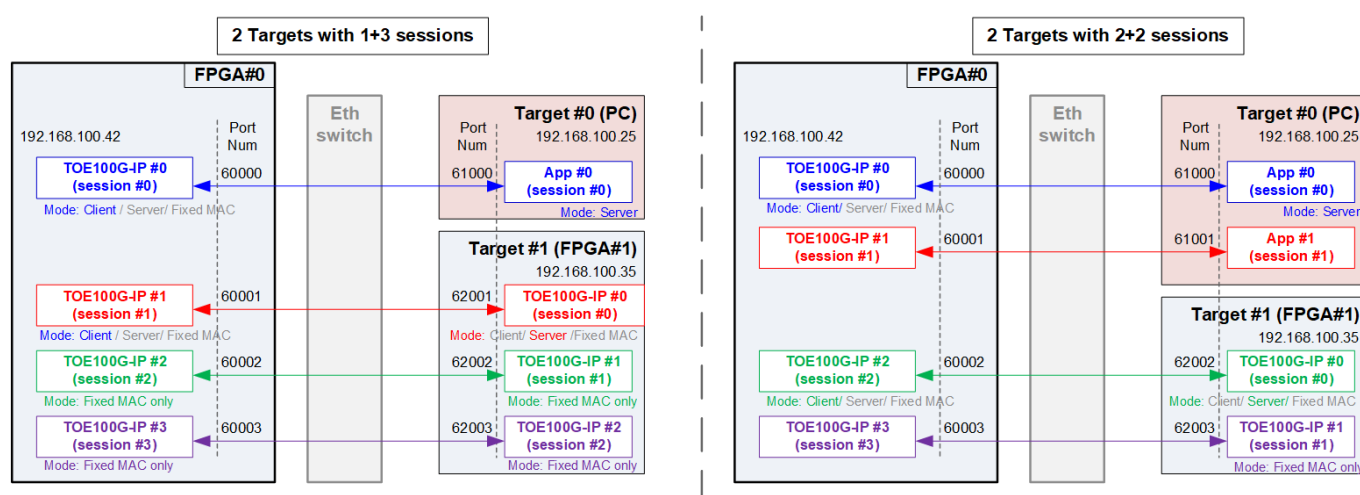


Figure 3-1 Example when running with two targets in test system

Figure 3-1 shows the example to run multisession design by using two target devices with four sessions. In the left side, the first target by PC runs one application for operating one session while the second target by FPGA#1 runs three TOE100G-IPs for operating three sessions. In the right side, the first target by PC runs two applications for operating two sessions while the second target by FPGA#1 runs two TOE100G-IPs for operating two sessions. In the real system, less than four sessions can be operated.

After FPGA (FPGA#0) boot-up, 100G Ethernet link up status (USER_RST_REG[16]) is polling. The CPU waits until link up is found. Next, welcome message is displayed and user selects the initialization mode of TOE100G-IP. There are three modes to be selection choices, i.e., Client, Server, and Fixed-MAC. Only the first session of each target can be assigned the initialization mode. The other sessions of the same target are fixed to be Fixed MAC.

When the target device is PC, it is recommended to configured the initialization mode on FPGA#0 to be Client mode. The IP initialization is completed when PC returns ARP reply after receiving ARP request. When two FPGAs are communicated, it is free to be initialized by several settings for the first session, i.e., Client <-> Server, Client <-> Fixed-MAC, and Fixed-MAC <-> Fixed MAC.

After receiving the initialization mode from the user, the default parameters in the selected mode are displayed on the console. The user selects to use default parameters or update the parameters before starting initialization. The example when the system is initialized in Client mode by using default parameters is shown in Figure 3-2.

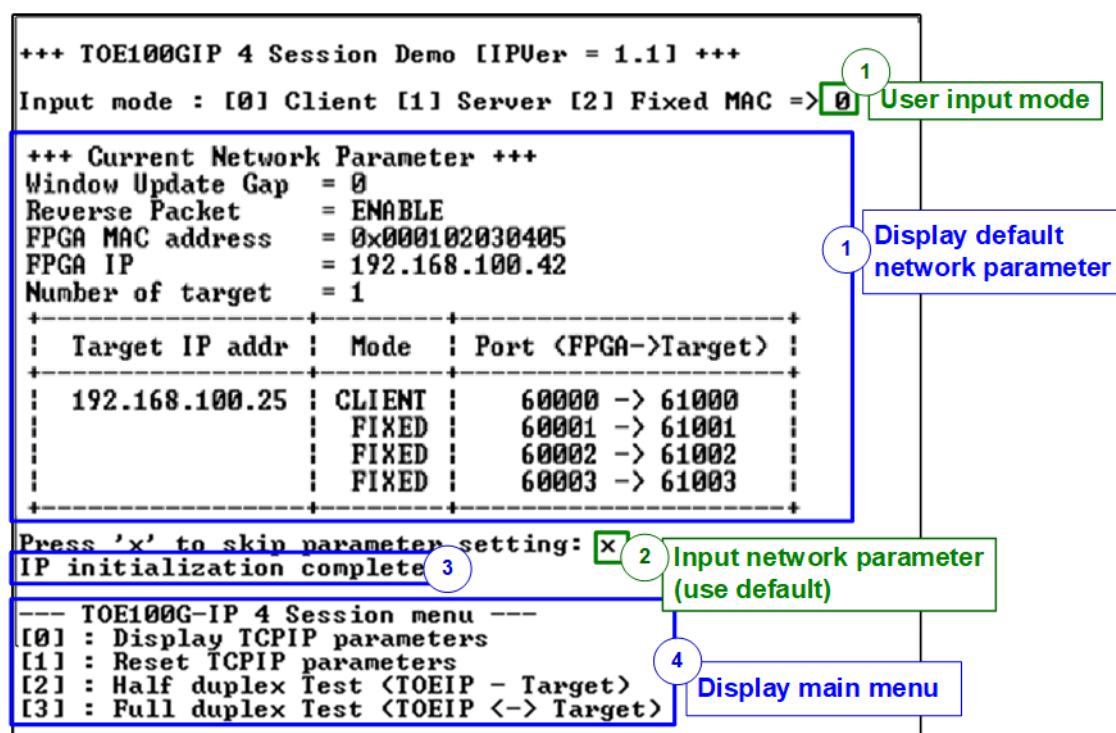


Figure 3-2 System initialization in Client mode by using default parameters

There are four steps to complete initialization sequence as follows.

- 1) CPU receives the initialization mode and then displays default parameters of the selected mode on the console.
- 2) User inputs 'x' to complete initialization process by using default parameters. Other keys are set for changing some parameters. More details for changing some parameters are described in Reset IP menu (topic 3.2).
- 3) CPU waits until all TOE100G-IPs finish initialization sequence (TOE_CMD_REG[0]='0').
- 4) Main menu is displayed. There are four test operations for user selection. More details of each menu are described as follows.

3.1 Display parameters

This menu is used to show current parameters of all active TOE100G-IPs, i.e., the initialization mode, reverse packet enable, Windows update threshold, source MAC address, destination IP address, source IP address, destination port, source port, and destination MAC address (when using Fixed-MAC mode). The sequence of display parameters is as follows.

- 1) Read common parameters that are shared for all sessions, i.e., FPGA MAC address, FPGA IP address, Window Update Gap, and reverse packet enable from each variable in firmware.
- 2) Print out each variable.
- 3) Read number of target devices and number of sessions for each target.
- 4) Read session parameters that can be set individually for each session, i.e., Initialization mode, Target IP address, Target MAC address (when using Fixed-MAC mode), FPGA port number, and Target port number from each variable in firmware.
- 5) Print out each variable on the table.

3.2 Reset IP

This menu is used to change TOE100G-IP parameters such as IP address and source port number. After setting new parameters to TOE100G-IP registers, the CPU resets the IP to re-initialize by using new parameters. Finally, the CPU monitors busy flag to wait until the initialization is completed. The sequence to reset IP is as follows.

- 1) Display current parameter value to the console.
- 2) Ask user to skip (use current parameters) or set new parameter values.
 - a. Press 'x' on keyboard to skip. The current parameters are used and continue to step 6.
 - b. Press other keys to start parameter value setting (continue to step 3).
- 3) Receive initialization mode from user and confirm that input is valid.
 - a. If input mode is invalid, mode value will not change and continue to step 4.
 - b. If input mode is valid and the value is changed from previous value, the current parameter set of new mode is displayed on the console. Next, user inputs 'x' to use the current parameters (continue to step 6). Otherwise, user inputs other keys to set the new parameters (continue to step 4).
- 4) Receive new common parameters from user, i.e., Window Update Gap, Reverse packet enable, FPGA MAC address, and FPGA IP address. If an input value is invalid, the parameter is not changed.
- 5) Received amount of target and parameters of each target from user, i.e., amount of session, Target MAC address (only in Fixed-MAC mode), Target IP address, and port number. If an input value is invalid, the input is not changed. However, if the amount of target is updated, user needs to set all parameters by valid value.
- 6) Force reset to all IPs by setting TOE_RST_REG[0]='1'.
- 7) For the sessions that are enabled,
 - i. Reset PattGen and PattVer logic by sending reset to user logic (USER_RST_REG[0]='1').
 - ii. Set all parameters to TOE100G-IP register such as TOE_SML_REG and TOE_DIP_REG.
- 7) For the first session of each target device,
 - i. De-assert TOE100G-IP reset by setting TOE_RST_REG[0]='0' to start IP initialization process.
 - ii. Wait until the TOE100G-IP completes initialization process (TOE_CMD_REG[0]='0').
 - iii. Read the Target MAC address (TOE_DML_REG and TOE_DMH_REG) to set the same value to the remaining session of the same target device.
- 8) For the remaining session of each target device,
 - i. Set Target MAC address (TOE_DML_REG and TOE_DMH_REG) by the same value as the first session.
 - ii. Set initialization mode (TOE_SRV_REG) to be Fixed-MAC mode to use the same Target MAC address with the first session.
 - iii. De-assert TOE_RST_REG[0] to '0' to start IP initialization.
 - iv. Wait until the TOE100G-IP completes initialization process (TOE_CMD_REG[0]='0').

3.3 Half Duplex Test

This menu is designed to transfer data in single direction for the initialized session. The user sets transfer mode to be send data, receive data, or no operation to each session individually. Next, transfer size and connection mode (active open for client mode or passive open for server mode) are assigned by user for the session that is set to send data or receive data. While the last parameter of send data command is packet size, the last parameters of receive data command is data verification flag (enable or disable). The operation is cancelled if some inputs are invalid.

To run the test, 32-bit incremental data is generated to send or verify data. The operation is finished when total data of all test session are transferred. When running the test with PC, test application (tcpdatatest) must be run.

The sequence of the test is as follows.

- 1) Display target IP address and port number of the first session.
- 2) Receive transfer mode, transfer size, packet size/data verification mode, and connection mode from user and verify if all inputs are valid.
- 3) Repeat step 2) to get the parameters of the next active session until the current session is the final active session.
- 4) Set UserReg registers following the transfer direction.
 - a. To send data, set transfer size (USER_TXLEN_REG), reset flag to clear initial value of PattGen (USER_RST_REG[0]='1'), and command register to start PattGen operation (USER_CMD_REG=0). After that, PattGen starts sending data to TOE100G-IP.
 - b. To receive data, set reset flag to clear initial value of PattVer (USER_RST_REG[0]='1') and command register to start PattVer operation with or without data verification (USER_CMD_REG=1/3). After that, PattVer starts verifying data from TOE100G-IP.
- 5) Display recommended parameters of test application on PC following connection mode.
 - a. For active connection mode, the parameters for running test application on PC by server mode are displayed. After that, "Press any key to proceed" is displayed to wait until user runs the test application on PC completely and enters some keys to continue the next step.
 - b. For passive connection mode, the parameters for running test application on PC by client mode are displayed.
- 6) Start open connection following connection mode setting.
 - a. For client mode (active open), CPU sets TOE_CMD_REG=2 (Open port) and set current state variable to WAIT_CONN.
 - b. For server mode (passive open), set current state variable to WAIT_CONN.
- 7) Wait until ConnOn status (USER_CMD_REG[2]) is asserted to '1'. After that, the current state variable changes to CONNECTED. After that, the data starts transferring.

8) The steps to run depends on transfer direction.

Send data command

- i. Confirm the state variable is equal to CONNECTED and TOE100G-IP is not busy (TOE_CMD_REG[0]='0'). If not, skip to operate the next session.
- ii. Continue the next step only when the connection is still ON (USER_CMD_REG[2]='1'). Otherwise, the state variable is set to ERROR.
- iii. Check total remaining length variable.
 - o If remaining length is equal to 0, run active close command by setting TOE_CMD_REG=3. Also, busy status of user logic (USER_CMD_REG[0]) must be equal to 0. After that, the state variable is set to CLOSE. If busy status is still asserted to '1', the state variable is set to ERROR.
 - o If remaining length variable is not equal to 0, run IP send operation by setting Packet size (TOE_PKL_REG), total transfer size (TOE_TDL_REG), and Send command (TOE_CMD_REG=0) respectively. After that, decrease total remaining length by TOE_TDL_REG value.

Note: TOE_TDL_REG is set to maximum transfer size (0xFFFF_FFC0) for each round except the final loop that is set to the total remaining length variable.

Receive data command

- i. Confirm the state variable is equal to CONNECTED.
- ii. Read the connection status. If connection status is OFF (USER_CMD_REG[2]='0'), set the state variable to CLOSED.

Every second the test progress and connection status of all sessions is displayed. If the data is still transferred, total amount of transmitted data (USER_TXLEN_REG) and received data (USER_RXLEN_REG) are read and displayed on the console.

- 9) For the session that has just run receive test, compare receive length of user logic (USER_RXLEN_REG) with the set value from user and read verification result (USER_CMD_REG[1]). Error message is displayed if receive length or data verification is error.
- 10) Calculate performance and show test result on the console.

3.4 Full duplex test

This menu is designed to run full duplex test by transferring data between FPGA and another device (PC/FPGA) in both directions by using the same port number at the same time. Five inputs are received from user, i.e., transfer mode (full-duplex or no operation), transfer size of both directions, packet size for send test, data verification mode for receive test, and connection mode (active open/close for client mode or passive open/close for server mode).

When running the test by using PC, the transfer size set on FPGA must be matched to the size set on test application (tcp_client_txrx_single). Connection mode on FPGA when running with PC must be set to passive (server operation).

The sequence of this test is as follows.

- 1) Display target IP address and port number of the first session.
- 2) Receive transfer mode, transfer size, packet size, data verification mode, and connection mode from the user and verify that all inputs are valid.
- 3) Repeat step 2) to get the parameters of the next active session until the current session is the final active session.
- 4) For the session that is set to full-duplex and the connection is passive mode, display the recommended parameters of test application run on PC from the current system parameters.
- 5) For the active session, set UserReg registers, i.e., transfer size (USER_TXLEN_REG), reset flag to clear the initial value of test pattern (USER_RST_REG[0]='1'), and command register to start PattGen and PattVer with data verification mode (USER_CMD_REG=0 or 2).
- 6) If there are some sessions setting connection mode to be active mode, "Press any key to proceed" is displayed to wait until user prepares the target device for connection creating. After that, user enters the key to continue the next step.
- 7) Start open connection following connection mode setting, similar to step 6) – 7) of Half duplex test.
- 8) Start data transferring by running following steps.
 - i. Confirm the state variable is equal to CONNECTED and TOE100G-IP is not busy (TOE_CMD_REG[0]='0'). If TOE100G-IP is busy, skip to operation the next session.
 - ii. Skip to the next step when the connection is still ON (USER_CMD_REG[2]='1'). Otherwise, confirm the connection mode is passive and remaining transmit length variable is equal to 0 before setting the state variable to CLOSED. If the session is active or the remaining transmit length is not equal to 0, the state variable changes to ERROR.
 - iii. Skip to the next step if the remaining transmit length variable is equal to 0. Otherwise, run IP send operation by setting Packet size (TOE_PKL_REG), total transfer size (TOE_TDL_REG), and Send command (TOE_CMD_REG=0) respectively. After that, decrease total remaining length by TOE_TDL_REG value.
Note: TOE_TDL_REG is set to maximum transfer size (0xFFFF_FFC0) which is aligned to packet size for each round except the final loop that is set to the total remaining length variable.
 - iv. If total amount of received data is equal to the set value and the connection mode is active, run active close command by setting TOE_CMD_REG=3. Also, busy status of user logic (USER_CMD_REG[0]) must be equal to 0 and set the stable variable to CLOSED. If busy is still asserted to '1', the state variable is set to ERROR.

Every second the test progress and connection status of all sessions is displayed. If the data is still transferred, total amount of transmitted data (USER_TXLEN_REG) and received data (USER_RXLEN_REG) are read and displayed on the console.

- 9) Check the result and the error, similar to step 9) of Half duplex test.
- 10) Calculate performance and show the test result on the console.

3.5 Function list in User application

This topic describes the function list to run TOE100G-IP operation.

unsigned int cal_strlen(unsigned int num)	
Parameters	num: integer input to calculate the string length
Return value	ret: the length of string for displaying
Description	Receive the input and then calculate the length of string to display this value in integer style.

void get_toeindex(unsigned int *num_target, unsigned int *num_session)	
Parameters	*num_target: pointer of the number of target variable *num_session: pointer of array that stores the number of sessions of each target
Return value	None
Description	Read all target IP address variable in firmware and then calculate the number of target and the number of sessions of each target. After that, return the result to num_target and num_session.

void init_param(void)	
Parameters	None
Return value	None
Description	This function is called to set the parameters and reset the IP, following described in topic 3.2.

void input_param(void)	
Parameters	None
Return value	None
Description	Receive test parameters from user for both common parameters (Reverse packet enable, Window threshold, FPGA MAC address, and FPGA IP address) and session parameters (Initialization mode, FPGA port number, Target IP address, Target port number, and Target MAC address when running in Fixed MAC mode). After receiving all parameters, the current value of all parameter is displayed.

unsigned int read_conon(void)	
Parameters	None
Return value	0: Connection is OFF, 1: Connection is ON.
Description	Read value from USER_CMD_REG register and return only bit2 value to show connection status.

void show_cursize(unsigned int *test_mode, unsigned int *cur_state)	
Parameters	test_mode: pointer of array that stores test_mode which can be set to No test (0), Send test (1), Receive test (2), or Full-duplex (3) cur_state: pointer of array that stores current state variable
Return value	None
Description	Read and display connection status of all session. If the connection is active, read current amount of transmitted data and received data from USER_TXLEN_REG and USER_RXLEN_REG and then display on the console in Byte, KByte, or MByte unit.

void show_ipaddr(unsigned int ip_addr)	
Parameters	ip_addr: IP Address in hexadecimal unit
Return value	None
Description	Display IP Address in decimal unit

void show_perf_header(unsigned int *test_mode)	
Parameters	test_mode: pointer of array that stores test_mode which can be set to No test (0), Send test (1), Receive test (2), or Full-duplex (3)
Return value	None
Description	When test mode is not no test (0), read Target IP address, Target port number, and FPGA port number to display as the header of the current status table.

void show_perf_line(unsigned int *test_mode)	
Parameters	test_mode: pointer of array that stores test_mode which can be set to No test (0), Send test (1), Receive test (2), or Full-duplex (3)
Return value	None
Description	When test mode is not no test (0), display straight line to be a part of the current status table.

void show_param(void)	
Parameters	None
Return value	None
Description	Display the current test parameters which consist of common parameters (Windows threshold, Reverse packet enable, FPGA MAC address, and FPGA IP address) and session parameters (Target MAC address, Target IP address, initialization mode, and port number). Total amount of target is also displayed.

void show_result(unsigned int *test_mode, unsigned int *cur_state, unsigned int *tot_rcv, unsigned int *total_len, unsigned int *err_rcv_ver)	
Parameters	test_mode: pointer of array that stores test_mode which can be set to No test (0), Send test (1), Receive test (2), or Full-duplex (3) cur_state: pointer of array that stores current state variable tot_rcv: pointer of array that stores total amount of received data total_len: pointer of array that stores total transferred data size err_rcv_ver: pointer of array that indicates data verification failed
Return value	None
Description	Display error message of receive operation if some error is found. After that, read USER_TXLEN_REG and USER_RXLEN_REG to display total amount of transmitted data and received data. Also, read the global parameters of timers to calculate total time usage to display in usec, msec, or sec unit. Finally, transfer performance is calculated and displayed in MB/s unit.

int toe_full_test(void)	
Parameters	None
Return value	0: The operation is successful -1: Receive invalid input or error is found
Description	Run Full duplex test following described in topic 3.4.

int toe_half_test(void)	
Parameters	None
Return value	0: The operation is successful -1: Receive invalid input or error is found
Description	Run half duplex data test following description in topic 3.3

void toe_tx_send(unsigned int toe_index, unsigned int *total_len, unsigned int *pac_size, unsigned int *round_size)	
Parameters	toe_index: the current number of TOE100G-IP for operating total_len: pointer of the total transfer size to send data pac_size: pointer of the packet size to send data round_size: pointer of the maximum transmit size for each test round
Return value	0: The operation is successful -1: Receive invalid input or error is found
Description	Set TOE100G-IP register to run send command by setting TOE_PKL_REG, TOE_TDL_REG, and TOE_CMD_REG.

void wait_ethlink(void)	
Parameters	None
Return value	None
Description	Read USER_RST_REG[16] and wait until ethernet connection is linked up

4 Test Software on PC

4.1 “tcpdatatest” for half duplex test



```

Command Prompt

D:\Temp>tcpdatatest.exe

[ERROR] The application requires at least 6 input parameters.

*****
TCP Data Transfer Test Version 1.2
*****
tcpdatatest.exe [Mode] [Dir] [ServerIP] [ServerPort] [ByteLen] [Pattern] [Window Scale]

[Mode]      PC Operation mode
             c:Client mode      s:Server mode
[Dir]       Transfer direction of PC
             t:Transmit data    r:Receive data
[ServerIP]  Server IP Address
[ServerPort] Server Port number(0-65535)
[ByteLen]   Transfer length(Byte)
[Pattern]   Disable/Enable Data Pattern in transferring
             0:Disable         1:Enable
[Window Scale] increase window size(optional)
             1:64K            2:128K            3:256K

[Example] tcpdatatest.exe s t 192.168.7.25 4000 34359738368 0

D:\Temp>_

```

Figure 4-1 “tcpdatatest” application usage

“tcpdatatest” is designed to run on PC for sending or receiving TCP data via Ethernet as server or client mode. PC of this demo should run in client mode. User sets parameters to select transfer direction and the mode. Six parameters are required as follows.

- 1) Mode: c –PC runs in Client mode and FPGA runs in Server mode
- 2) Dir: t – transmit mode (PC sends data to FPGA)
r – receive mode (PC receives data from FPGA)
- 3) ServerIP: IP address of FPGA when PC runs in Client mode (default is 192.168.7.42)
- 4) ServerPort: Port number of FPGA when PC runs in Client mode (default is 4000)
- 5) ByteLen: Total transfer size in byte unit. This input is used in transmit mode only and ignored in receive mode. In receive mode, the application is closed when the connection is terminated. In transmit mode, ByteLen must be equal to the total transfer size on FPGA that is set in receive data test menu.
- 6) Pattern:
 - 0 – Generate dummy data in transmit mode or disable data verification in receive mode.
 - 1 – Generate incremental data in transmit mode or enable data verification in receive mode.

Note: Window Scale: Optional parameter which is not used in the demo.

Transmit data mode

Following sequence is the sequence when test application runs in transmit mode.

- 1) Get parameters from the user and verify that all inputs are valid.
- 2) Create the socket and set socket options.
- 3) Create the new connection by using server IP address and server port number.
- 4) Allocate 1 MB memory to be send buffer.
- 5) Skip this step if the dummy pattern is selected. Otherwise, generate the incremental test pattern to send buffer.
- 6) Send data out and read total sent data from the function.
- 7) Calculate remaining transfer size.
- 8) Print total transfer size every second.
- 9) Repeat step 5) – 8) until the remaining transfer size is 0.
- 10) Calculate total performance and print the result on the console.
- 11) Close the socket and free the memory.

Receive data mode

Following sequence is the sequence when test application runs in receive mode.

- 1) Follow the step 1) – 3) of Transmit data mode.
- 2) Allocate memory to be receive buffer.
- 3) Read data from the receive buffer and increase total amount of received data.
- 4) This step is skipped if data verification is disabled. Otherwise, received data is verified by the incremental pattern. Error message is printed out when data is not correct.
- 5) Print total amount of received data every second.
- 6) Repeat step 3) – 5) until the connection is closed.
- 7) Calculate total performance and print the result on the console.
- 8) Close socket and free the memory.

4.2 “tcp_client_txrx_single” for full duplex test



```

Command Prompt
D:\Temp>tcp_client_txrx_single.exe

*****
TCP Tx Rx Version 1.0
*****
tcp_client_txrx_single.exe [ServerIP] [ServerPort] [ByteLen] [Verification]

[ServerIP]      Server IP Address
[ServerPort]    Server Port number(0-65535)
[ByteLen]       Transfer length(Byte)
[Verification]  Disable/Enable Verification in transferring
                0:Disable          1:Enable

[Example]  tcp_client_txrx_single.exe 192.168.40.42 60000 137438953440 0

D:\Temp>

```

Figure 4-2 “tcp_client_txrx_single” application usage

This application is similar to “tcp_client_txrx_40G” application but it runs for one round, not forever loop. “tcp_client_txrx_single” application is designed to run on PC for sending and receiving TCP data through Ethernet by using the same port number at the same time. The application is run in Client mode, so user needs to input server parameters (the network parameters of TOE100G-IP). As shown in Figure 4-2, there are four parameters to run the application, described as follows.

- 1) ServerIP : IP address of FPGA
- 2) ServerPort : Port number of FPGA
- 3) ByteLen : Total transfer size in byte unit. This is total size to transmitted data and received data.
- 4) Verification :
 - 0 – Generate dummy data for sending function and disable data verification for receiving function. This mode is used to check the best performance of full-duplex transfer.
 - 1 – Generate incremental data for sending function and enable data verification for receiving function.

The sequence of test application is as follows.

- (1) Get parameters from the user and verify that the input is valid.
- (2) Create the socket and set socket options.
- (3) Create the new connection by using server IP address and server port number.
- (4) Allocate 64 KB memory for send buffer and receive buffer.
- (5) Generate incremental test pattern to send buffer when the test pattern is enabled. Skip this step if dummy pattern is selected.
- (6) Send data out, read total amount of transmitted data from the function, and calculate remaining send size.
- (7) Read data from the receive buffer and increase total amount of received data.
- (8) Skip this step if data verification is disabled. Otherwise, data is verified by incremental pattern. Error message is printed out when data is not correct.
- (9) Print total amount of transmitted data and received data every second.
- (10) Repeat step 5) – 9) until total amount of transmitted data and received data are equal to ByteLen, set by user.
- (11) Calculate performance and print the result on the console.
- (12) Close the socket.

5 Revision History

Revision	Date	Description
1.0	11-Mar-22	Initial version release
1.1	21-Mar-22	Correct Figure 2-1