Ultra-Low Latency 10G Ethernet IP Solution

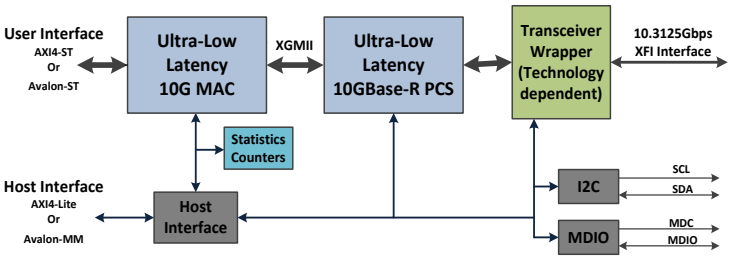
Product Brief (HTK-ULL10G-ETH-32-FPGA)

The 10Gbps 32-bit Ethernet IP solution offers a fully integrated IEEE802.3-2015 compliant package for NIC (Network Interface Card) and Ethernet switching applications. This industry leading ultra-low latency solution is specifically targeted for demanding financial, high frequency trading and HPC applications.

- ***Round Trip Latency of 52.7ns + Device Specific Transceiver Latency***

As shown in the figure below, the 10Gbps Ethernet IP solution includes:

- ***Ultra-Low latency MAC; Tx = 12.4ns , Rx = 15.5ns;*** (32-bit user interface mode with FCS generation and checking)
- ***Ultra-Low latency PCS; Tx = 12.4ns , Rx = 12.4ns;***
- Technology dependent transceiver wrapper for Altera and/or Xilinx FPGAs
- Statistics counter block (for RMON and MIB)
- MDIO and I2C cores for external module and optical module status/control



A complete reference design using a L2 (MAC level) packet generator/checker is also included to facilitate quick integration of the Ethernet IP in a user design. A GUI application interacts with the reference design's hardware elements through a UART interface (a PCIe option is also available). An application (with optional basic Linux PCIe driver/API) is also provided for memory mapped read/write access to the internal registers. See **Appendix A** for details.

MAC and PCS cores are designed with 32-bit data path to take advantage of high performance fabrics of the 28nm and 20nm FPGAs. This implementation approach also delivers industry's lowest latency and low area footprint.

As the PCS and transceiver wrapper is included with the Ethernet IP solution, the line side directly connects the 10.3125Gbps FPGA transceiver to the optical module (SFP+, XFP etc).

Ethernet IP solution implements two user (application) side interfaces. The register configuration and control port can be 32-bit AXI4-Lite or Avalon-MM interface. Depending upon the application layer, user can select a

32-bit or 64-bit AXI-4 Streaming or Avalon Streaming bus to interface with the MAC block.

10Gbps Ethernet IP supports advanced features like per-priority pause frames (compliant with 802.3bd specifications) to enable Converged Enhanced Ethernet (CEE) applications like data center bridging that employ IEEE 802.1Qbb Priority Flow Control (PFC) to pause traffic based on the priority levels.

Features Overview

MAC Core Features

- Implements the full 802.3 specification with preamble/SFD generation, frame padding generation, CRC generation and checking on transmit and receive respectively.
- Implements 802.3bd specification with ability to generate and recognize PFC pause frames.
- Implements reconciliation sublayer functionality with start and terminate control characters alignment, error control character and fault sequence insertion and detection.
- PCS layer XGMII interface implemented as 32-bit (double data rate) DDR interface to operate at 10Gbps for direct interface to 10GBase-R, XAUI and RXUAI cores
- Deficit Idle Count (DIC) mechanism to ensure data rates of 10Gbps at the transmit interface.
- Optional padding of frames if the size of frame is less than 64 bytes.
- Implements fully automated XON and XOFF Pause Frame (802.3 Annex 31A) generation and termination providing flow control without user application intervention. Non PFC Mode only with use of support wrapper containing Rx User Interface Buffer.
- Pause frame generation additionally controllable by user application offering flexible traffic flow control.
- Support for VLAN tagged frames according to IEEE 802.1Q.
- Support any type of Ethernet Frames such as SNAP / LLC, Ethernet II/DIX or IP traffic.
- Discards frames with mismatching destination address on receive (except Broadcast and Multicast frames).
- Programmable Promiscuous mode support to omit MAC destination address checking on receive path.
- Optional multicast address filtering with 64-bit Hash Filtering table providing imperfect filtering to reduce load on higher layers.
- High speed CRC-32 generation and checking.

- Optional prevention of CRC appending in frame data by MAC to allow CRC to be pre-embedded in frame data by user application.
- Optional insertion of error control character in transmitted frame data.
- Optional forwarding of the CRC field to user application interface.
- Programmable frame maximum length providing support for any standard or proprietary frame length (e.g. 9K-Bytes Jumbo Frames).
- Status signals available with each Frame on the user interface providing information such as frame length, VLAN frame type indication and error information.
- Optional padding termination on RX path for NIC applications or forwarding of unmodified data to the user interface.
- Optional internal XGMII Loop-back with use of support wrapper containing XGMII Loop-back Buffer.
- Statistics indicators for frame traffic as well as errors (alignment, CRC, length) and pause frames.
- Altera Avalon or Xilinx AXI4 interface compliant user interface with use of support wrapper containing Tx and Rx User Interface Buffers.
- Implements statistics and event signals providing support for 802.3 basic and mandatory managed objects as well as IETF Management Information Database (MIB) package (RFC 2665) and Remote Network Monitoring (RMON) required in SNMP environments.

PCS Core Features

- Implements 10GBase-R PCS core compliant with IEEE 802.3-2008 Specifications.
- Implements a 32-bit XGMII interface to operate at 10Gbps for 10G Ethernet.
- Implements 64b/66b encoding/decoding for transmit and receive PCS using 802.3-2008 specified control codes.
- Implements 10G scrambling/descrambling using 802.3-2008 specified polynomial $1 + x^{39} + x^{58}$.
- Implements 66-bit block synchronization state machine as specified in 802.3-2008 specifications.
- Automatic clock compensation without the need for Inter Packet Gap (IPG) insertion/deletion that is achieved with use of support wrapper containing Rx XGMII Buffer.
- Implements gear-box logic to connect to 32-bit transceivers for line side. The 32-bit interface operates at the transceiver reference clock.
- Implements Bit Error Rate (BER) monitor for monitoring excessive error ratio. In addition, the core implements various status and statistics required by the IEEE 802.3-2008 such as block synchronization status and test mode error counter.
- Implements optional XGMII remote loopback to loopback data received from Rx PCS back to Tx PCS with use of support wrapper containing XGMII remote loop-back Buffer.

Licensing and Maintenance

- ***NO yearly maintenance fees for upgrades and bug fixes***
- Basic core licensing for a single vendor (either Xilinx or Altera) compiled (synthesized netlist) binary
- Option for vendor and device family agnostic source code (Verilog) license

Contact and Sales Information

For further information, contact sales representative at:

Phone: +1-301-528-8074

Email: sales@hiteksys.com

Resource Utilization

The utilization summary of the 10G Ethernet solution is given in following tables. The utilization numbers are best in class as compared to other available 10G Ethernet cores with comparable feature set.

The Ethernet solution has been fully verified on different hardware platforms for both Altera and Xilinx FPGAs and has also been verified for interoperability with other 10G capable devices.

10G ULL Ethernet IP - Resource Usage for Xilinx Devices

Device	*User Interface (AXI4)*	*Priority Flow Control (PFC)*	*Slice LUTS*	*Slice Registers*	*BRAMs*
UltraScale/ UltraScale+	32-Bit	No	2,437	1,704	*18K = 0; 36K = 0*
		Yes	2,534	2,129	*18K = 0; 36K = 0*
7-Series	32-Bit	No	2,455	1,703	*18K = 0; 36K = 0*
		Yes	2,565	2,128	*18K = 0; 36K = 0*

Note:

- *Support wrapper with MAC and PCS Registers adds additional 529 Slice LUTs and 1276 Slice Registers.*
- *Register based RMON statistics block adds additional 1948 Slice LUTs and 1807 Slice Registers.*

10G ULL Ethernet IP - Resource Usage for Altera Devices

Device	*User Interface (Avalon)*	*Priority Flow Control (PFC)*	*COMB. ALUTs*	*Registers*	*Memory M20K*
Arria 10	32-Bit	No	2,346	1,708	0
		Yes	2,415	2,135	0
Stratix V	32-Bit	No	2,351	1,703	0
		Yes	2,428	2,128	0

Note:

- *Support wrapper with MAC and PCS Registers adds additional 596 Comb. ALUTs and 1093 Registers.*
- *Register based RMON statistics block adds additional 2003 Comb. ALUTs and 1808 registers.*

Performance (Tx And Rx Latency)

The performance of the 10G Ethernet solution is represented here in terms of individual latencies of transmit and receive paths, i.e. the time between the first bit of data input at 10G Ethernet MAC and the first bit of data output at PCS-Transceiver interface. These numbers will change with the change in programmable threshold levels used for reading the user interface FIFOs. For the latencies given here, the thresholds for both transmit and receive User FIFOs were set to minimum possible values for correct operation. Data path latency is also dependent upon the type of user interface FIFO used in the design. FIFO implementation can either be a SCFIFO (Single Clock FIFO, when the MAC and application clock are same) or can be a DCFIFO (Dual Clock FIFO, when the MAC and application clock are different). Following table lists the latencies for various user interface options.

Technology	*User Interface*	*MAC+PCS Latency (ns)*	
		Tx	*Rx*
Xilinx	**32-bit (NO FIFO)**	**24.8**	**27.9**
	32-bit (DCFIFO)	40.38	52.8
	64-bit (SCFIFO)	43.4	46.6
	64-bit (DCFIFO)	43.4	55.9
Altera	**32-bit (NO FIFO)**	**24.8**	**27.9**
	32-bit (DCFIFO)	40.38	55.9
	64-bit (SCFIFO)	43.4	49.7
	64-bit (DCFIFO)	43.4	59.0

Deliverables

- Ethernet wrapper design with:
  - Top level MAC and PCS wrappers (source files, Verilog) for user specific customizations
  - Compiled synthesizable binaries (Netlists) for MAC and PCS cores
  - Compiled synthesizable binaries (Netlists) for L2 packet generator and checker
  - Technology specific transceiver wrappers for the selected device family
  - Source code RTL (Verilog) for RMON and Register-File blocks
- UART/PCIe interface based reference design with:
  - Top level wrapper (source files, Verilog) for user specific customizations
  - Compiled synthesizable binaries (Netlists) for the I2C and MDIO cores
  - Compiled synthesizable binaries (Netlists) for the UART or PCIe host interface
- Encrypted RTL for MAC, PCS and packet generator/checker for simulation
- Constraint files and synthesis scripts for design compilation
- PCIe driver/API (source files, C) for Linux with the optional PCIe host interface
- GUI application (Linux only) for interfacing to the reference design
- Design guide(s) and user manual(s)

Ported/Validated Modules List

1. *HiTech Global HTG-K800, HTG-K816 and HTG-830*; ***Xilinx Virtex Ultrascale*** and ***Kintex Ultrascale*** FPGAs; Interface through FMC (HTG-FMC-X4SFP+) and Z-Ray (HTG-ZR-X3SFP+) SFP+ modules.
   (http://hitechglobal.com/Boards/Kintex-UltraScale.htm)
   (http://www.hitechglobal.com/Boards/Kintex_UltraScale_half-size_PCIe.htm)
   (http://hitechglobal.com/boards/Virtex-UltraScale-FPGA.htm)
2. *HiTech Global HTG-K700****; Xilinx Kintex-7*** FPGA; Interface through FMC SFP+ module (HTG-FMC-X4SFP+)
   (http://www.hitechglobal.com/Boards/Kintex-7_PCIE.htm)
3. *Xilinx VC707* (***Xilinx Virtex-7*** 485) and *Xilinx KC705* (***Xilinx Kintex-7*** 325) Evaluation modules
4. *HiTech Global HTG-510*; ***Altera Stratix-V*** FPGA, with integrated QSFP+ and SFP+ interfaces
   (http://hitechglobal.com/Boards/Stratix-V_PCIExpress.htm)
5. *Solarflare SFA7942Q ApplicationOnload Engine (AOE);* ***Altera Stratix-V*** FPGA, with integrated QSFP+ support
   (http://www.solarflare.com/applicationonload-engine)

A. Reference Design Details

A.1 Overview

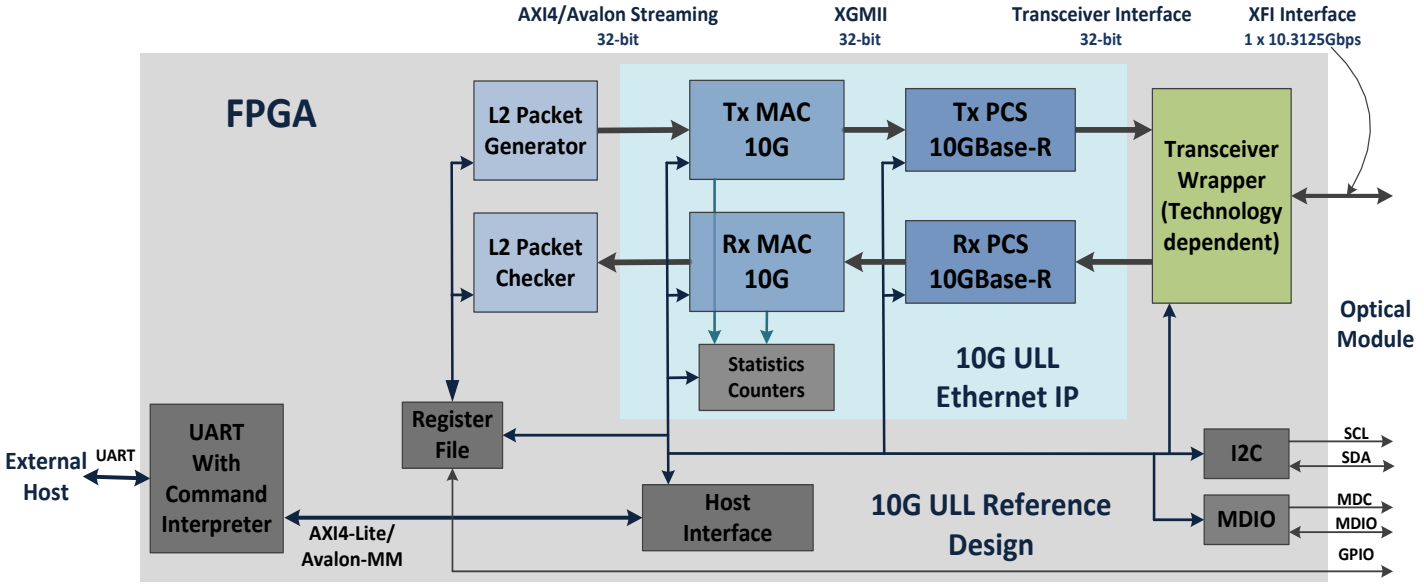
A 10Gbps reference design is included as part of the IP deliverable to facilitate quick L1 and L2 layer testing and verification of the 10Gbps Ethernet on target platform. The capability to run the L1 PRBS pattern and configure each transceiver independently can be for used for a fast module bring-up in the lab and can also be used for factory diagnostics.

The UART (normally through an onboard USB-to-UART converter chip) based 10G Ethernet reference design can be seamlessly ported to various COTS FPGA networking and evaluation modules (see section for the list of verified modules). A GUI application controls the register read/writes to the FPGA through a UART core with integrated command interpreter. Both Linux and Windows platforms are supported for the UART based interface control.

This reference design can also be used on custom embedded design where the FPGA connects to the host processor via a PCIe interface. For the PCIe control interface, GUI application is hosted on a Linux platform (as PCIe driver/API is provided for Linux OS only).

A.2 Functional Description

Following figure shows the connectivity and the elements of the 10Gbps Ethernet IP reference design. Usually the UART interface from the FPGA connects to an external (can be on the same module as well) USB-UART converter. A Linux or Windows host (through a USB port) running the GUI application is used to configure and control the 10G Ethernet. I2C and GPIO interfaces included in the reference design can be used to control any optical module on the target platform including the XFP+ and XFP compliant modules.



For L1 (physical layer verification and testing) GUI application provides an interface to independently control and configure 10.3125Gbps transceiver used for 10G Ethernet transport. User can configure the transceiver to run various PRBS pattern and configure various transceiver parameters like transmit voltage, transmit pre-emphasis, receive equalization and receive gain.

For L2 testing, GUI application uses the 10Gbps packet generator/checker inside the FPGA to generate and check MAC frames up to full line rate. Packet generator supports a basic rate control mechanism to control the packet/data rate on the interface. Generator can be configured for fixed size as well as pseudo random packet size packet transmission. An incrementing counter is used as payload for the MAC frames. Checker on the receive side verifies the payload of receive MAC frames and reports error in the payload.

A comprehensive set of transmit and receive counters in the MAC core provide a detailed view of the packet statistics including various error types.

Following is a snapshot for the GUI application for the L2 packet test results screen.