

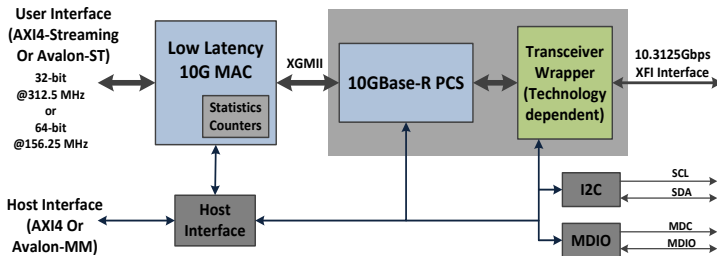
Low Latency 10G Ethernet IP Solution

Product Brief (HTK-LL10G-ETH-32-FPGA)



The 10Gbps 32-bit Ethernet IP solution offers a fully integrated IEEE802.3-2008 (802.3ae) compliant package for NIC (Network Interface Card) and Ethernet switching applications. This extremely low latency solution is specifically targeted for demanding financial, high frequency trading and HPC applications. As shown in the figure below, the 10Gbps Ethernet IP includes:

- **Lowest latency MAC; Tx = 41.6ns , Rx = 76.8ns;** (32-bit user interface mode)
- Flexible 10GBase-R PCS options with XFI interface for direct SFP+/XFP attachment
- 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 PCIe interface (a UART option is also available). A basic Linux PCIe driver/API is also provided for memory mapped read/write access to the internal registers. See **Appendix A** for details.

MAC core is designed with 32-bit data path operating at 312.5MHz to take advantage of high performance fabrics of the 45nm and 28nm FPGAs. This implementation approach also delivers industry's lowest latency and low area footprint.

Depending on the target vendor and device family, PCS layer can be a soft solution using 10.3125Gbps transceiver or an integrated Hard-IP (like in Altera's Stratix V device). PCS implementations are based upon 64-bit data path operating at 156.25MHz.

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 access port can either be a 32-bit AXI4 interface or a 32-bit Avalon-MM interface. IP solution provides a highly flexible 10Gbps traffic port interface options. Depending upon the application layer,

user can select an AXI-4 streaming bus or an Avalon Streaming bus to interface with the MAC block. In either mode, the MAC interface bus width is selectable as 32-bit @ 312.5MHz or 64-bit @ 156.25MHz.

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.
- Multiple user interface options for the MAC data path; AXI-4 streaming or Avalon Streaming; 32-bit data path 312.5MHz or 64-bit data path 156.25MHz
- PCS layer XGMII interface implemented as 64-bit (single data rate) SDR interface at 156.25MHz for direct interface to 10GBase-R, XAUI and RXAUI 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.
- 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.

Low Latency 10G Ethernet IP Solution

Product Brief (HTK-LL10G-ETH-32-FPGA)



- Optional multicast address filtering with 64-bit HASH Filtering table providing imperfect filtering to reduce load on higher layers.
- CRC-32 generation and checking at high speed using Galois field multipliers and alternate polynomials.
- 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.
- Statistics indicators for frame traffic as well as errors (alignment, CRC, length) and pause frames.
- Altera Avalon or Xilinx AXI4 interface compliant user (FIFO) interface.
- Transmit and Receive FIFOs with configurable depths having a default depth of 1KB/512B (128 64-bit/32-bit words) each, according to user interface bus width.
- 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.

See **Appendix B** for functional details of MAC core.

PCS Options

- Multiple 10GBase-R PCS options are supported by the 10G Ethernet solution, including;
 - Xilinx's 10GBase-R (free of cost) core; Default option for Xilinx; Nearly all PCS blocks are part of the Xilinx transceiver's Hard-IP
 - Hitek System's 10GBase-R core; Default option for Altera, except for Stratix V family
 - Altera's soft 10GBase-R core (require licensing from Altera)
 - Altera's integrated 10GBase-R hard-IP (Stratix V only)
- For devices without the support for the 10.3125Gbps transceivers, Ethernet IP solution can be configured to operate with vendor specific XAUI and RXAUI cores.

Licensing and Maintenance

- ***True sign once licensing with NO yearly maintenance fees***
- Basic core licensing for a single vendor (either Xilinx or Altera) compiled (synthesized) binary
- Additional vendor license provided at only 50% cost of the base license. This allows for cost effective multi-vendor designs with identical user and control interfaces.
- Other licensing options include:
 - Vendor and device family agnostic source code (Verilog) license
 - A ***low cost board locked*** license for low budget prototyping (upgradeable to full license)

Contact and Sales Information

Phone: +1-301-528-2244

Email: info@mantaro.com

Low Latency 10G Ethernet IP Solution

Product Brief (HTK-LL10G-ETH-32-FPGA)



MAC Resource Utilization

The MAC core 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. Total Ethernet utilization is dependent upon PCS core and device selection. In the best case scenario, with integrated 10GBase-R hard IP (in Altera's Stratix V series and Xilinx Virtex-6, Virtex-7, Kintex-7 series FPGAs), total resource utilization is approximately equal to MAC core utilization.

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 MAC - Resource Usage for Xilinx Devices

Device	User Interface Width	RMON and MDIO	Slices	Slice LUTs	Slice Registers	BRAMs
Virtex-6 (-2C Speed)	AXI-4, 32-Bit	No	1,042	3,316	3,034	2
		Yes	2,106	4,477	4,585	2
	AXI-4, 64-Bit	No	1,069	3,438	3,289	2
		Yes	2,152	4,587	4,843	2
	Avalon, 32-Bit	No	1,037	3,267	3,033	2
		Yes	1,871	4,692	4,655	2
Avalon, 64-Bit	No	1,107	3,413	3,295	2	
	Yes	2,173	4,697	4,910	2	
Virtex-5 (-2C Speed)	AXI-4, 32-Bit	No	1,507	3,467	3,039	2
		Yes	2,114	5,317	4,653	2
	AXI-4, 64-Bit	No	1,640	3,551	3,297	2
		Yes	2,252	5,401	4,910	2
	Avalon, 32-Bit	No	1,537	3,462	3,038	2
		Yes	2,150	5,312	4,652	2
Avalon, 64-Bit	No	1,608	3,540	3,293	2	
	Yes	2,248	5,390	4,906	2	

10G MAC - Resource Usage for Altera Devices

Device	User Interface Width	RMON and MDIO	COMB. ALUTs	Memory ALUTs	Registers	Memory M9K
Stratix-IV (-C2 Speed)	Avalon, 32-Bit	No	2,767	69	2,862	7
		Yes	3,773	69	4,377	7
	Avalon, 64-Bit	No	2,797	69	3,029	9
		Yes	3,798	69	4,545	9

MAC Performance (Latency)

The performance of the Ethernet MAC core is represented here in terms of individual latencies of transmit and receive paths, i.e. the time between the first bit of data input and the first bit of data output. This 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 1 and testing was done using 64 bytes of frame data. 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.

Low Latency 10G Ethernet IP Solution

Product Brief (HTK-LL10G-ETH-32-FPGA)



<i>Technology</i>	<i>User Interface</i>	<i>Latency (ns)</i>	
		<i>Tx</i>	<i>Rx</i>
Xilinx (Virtex-6)	32-bit (SCFIFO)	41.6	76.8
	32-bit (DCFIFO)	57.6	83.2
	64-bit (SCFIFO)	44.8	83.2
	64-bit (DCFIFO)	64	115.2
Altera (Stratix IV)	32-bit (SCFIFO)	41.6	76.8
	32-bit (DCFIFO)	48	83.2
	64-bit (SCFIFO)	44.8	83.2
	64-bit (DCFIFO)	51.2	115.2

Deliverables

- Compiled synthesizable binaries or encrypted RTL for the MAC core
- Source code RTL (Verilog) for I2C, MDIO, RMON and Register-File blocks
- Self checking behavioral models and test benches for simulation
- Constraint files and synthesis scripts for design compilation
- A complete PCIe/UART host interface based reference design with:
 - Top level wrapper (source files, Verilog) for user specific customizations
 - Source files (Verilog) for the PCIe application layer
 - Binaries for a basic L2 packet generator and checker
 - PCIe driver/API (source files, C) for Linux
 - UART and command interpreter blocks with the optional UART host interface
 - GUI application (Linux only for PCIe, Linux and Windows for UART) for interfacing to the reference design
- Design guide(s) and user manuals
- USA based technical support by developers

A. Reference Design Details

A.1 Overview

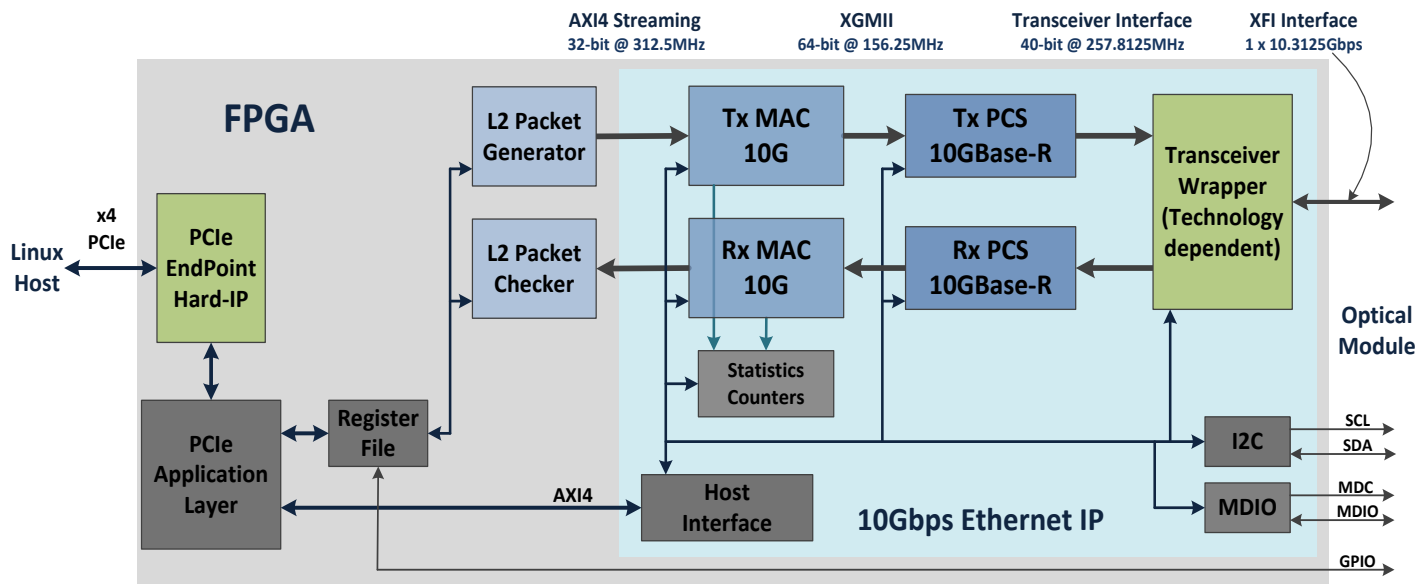
A 10Gbps reference design is included as part of the IP deliverable to facilitate quick L1 and L2 layer testing and verification of the 10Gbps Ethernet on target platform. The capability to run the L1 PRBS pattern and configure each transceiver independently can be for used for a fast module bring-up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

The PCIe based 10Gbps Ethernet reference design can be seamlessly ported to various COTS PCIe from factor FPGA networking and evaluation modules (see section for the list of verified modules). 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.

For the embedded designs and evaluation modules that do not have PCIe control interface but do have UART debug interface (usually through a USB-to-UART converter chip), 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.

A.2 Functional Description

Following figure shows the connectivity and the elements of the 10Gbps Ethernet IP reference design. A Linux host (embedded or standard PC) running a 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.

Low Latency 10G Ethernet IP Solution

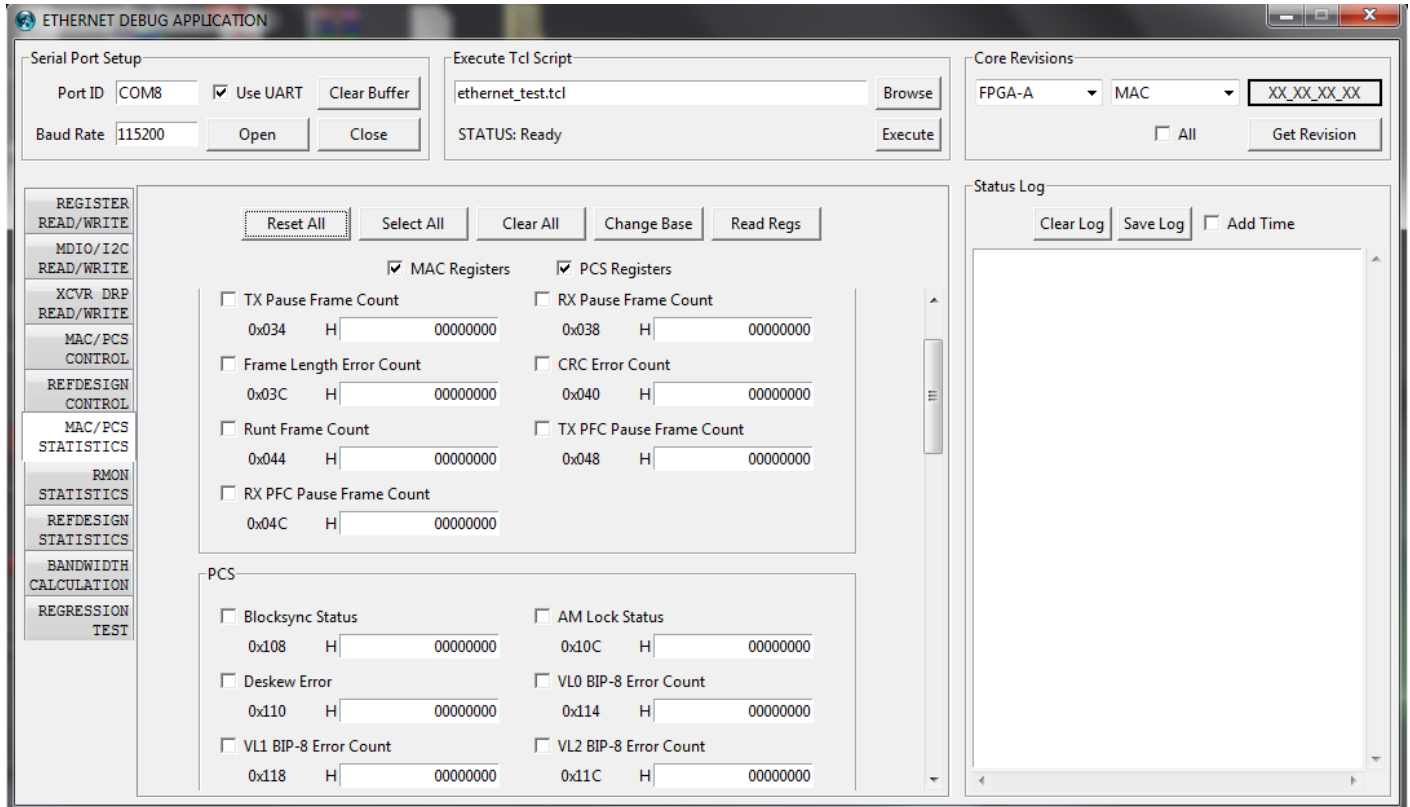
Product Brief (HTK-LL10G-ETH-32-FPGA)



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.

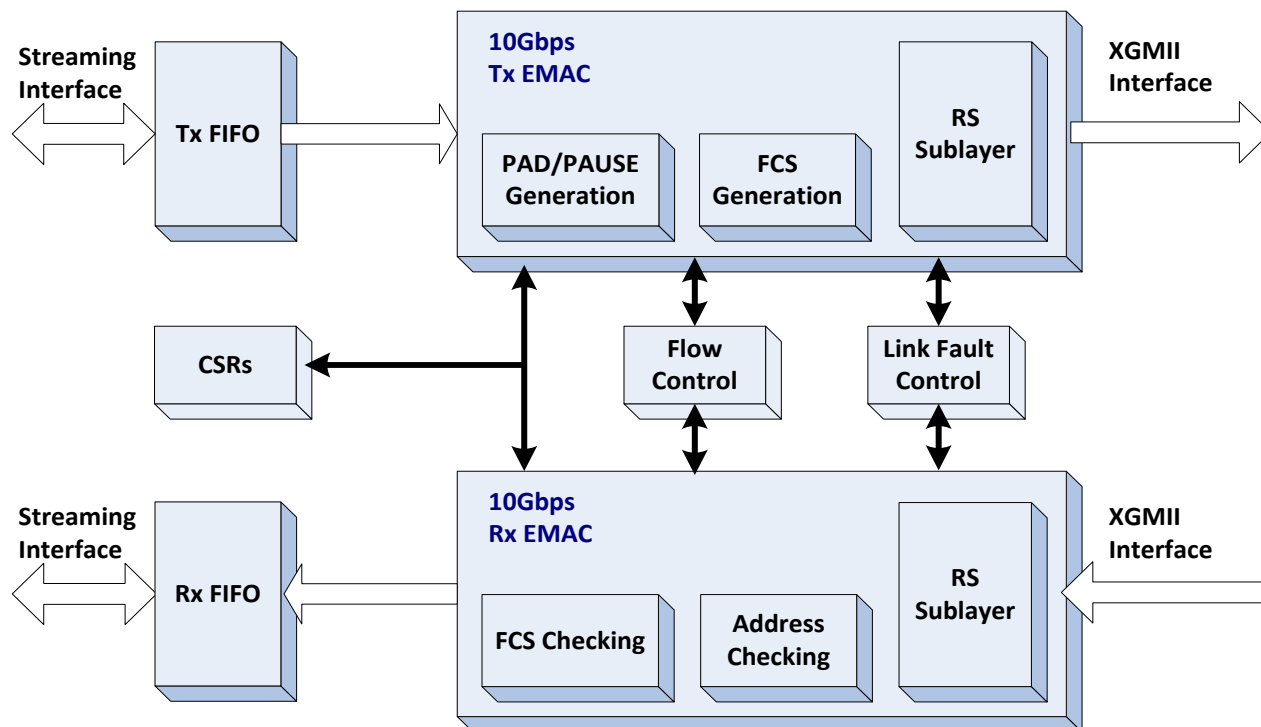


A.3 Ported/Validated Modules List

1. HTG-V6HXT-x16PCIE; Virtex-6 HXT FPGA, x16 PCIe Module with 2 QSFP+ and 2 SFP+ interfaces (http://www.mantaro.com/products/development_platforms/Xilinx_V6HXT_x16PCIE.htm)
2. HTG-V6HXT-x8PCIE; Virtex-6 HXT FPGA, x8 PCIe Module with 4 SFP+ interfaces (http://www.mantaro.com/products/development_platforms/Xilinx_V6HXT_x8PCIE.htm)
3. HTG-S4GT-PCIE; Virtex-6 HXT FPGA, x8 PCIe Module with 2 SFP+ interfaces (http://www.mantaro.com/products/development_platforms/Altera_S4G_PCIE.htm)

B. MAC Functional Overview

The following figure shows the architecture for the 10 Gigabit Ethernet MAC.



B.1 Reconciliation Sublayer (RS) Operation

The RS layer is responsible to map the data to/from the MAC sublayer to the XGMII interface. The RS layer provides a 32-bit interface with data sampled at positive edge of the 312.5 Mhz XGMII clock. This data interface is then converted to 64-bit interface (using the XGMII FIFOs) with data sampled at positive edge of the 156.25 XGMII clock. The 64-bit data is organized into eight 8-bit lanes with a control bit available for each lane.

B.2 MAC Sublayer Operation

The MAC sublayer is responsible to perform transmit and receive operations. The transmit MAC block transmits frames from a user application interface to the reconciliation sublayer, which then transmits these frames to the XGMII physical interface. The receive MAC block receives Ethernet frames from the reconciliation sublayer, validates the Ethernet frame and transfers this frame to the user application interface. The following description defines the various functions performed by transmit and receive Ethernet MAC engines.

Transmit Ethernet MAC

The transmit Ethernet MAC performs the following main functions:

- Accepts data including Destination Address, Source Address and length field from the MAC client.
- Appends preamble and SFD to the Ethernet frames.

Low Latency 10G Ethernet IP Solution

Product Brief (HTK-LL10G-ETH-32-FPGA)



- Inserts PAD field for frames with length less than minimum frame length (64 bytes) and runt frames are not configured to pass through.
- Calculates and Appends proper FCS (CRC-32) value to outgoing frames (unless CRC is pre-embedded in frame) and verifies full octet boundary alignment.
- Delays transmission of frame data for specified inter-frame gap period.
- Controls Inter-frame gap timing for LAN mode of operation.
- Generates preamble and SFD field before frame transmission.
- Manages local device flow control by generating PAUSE control frames.
- Manages Remote device congestion by transiting to HALT state for a specified time quanta.

Receive Ethernet MAC

The receive Ethernet MAC performs the following main functions:

- Receives a frame from the RS sub layer via a 32-bit data bus.
- Presents to the MAC client sublayer frames that are either frames with group address or directly addressed to the local station (Address recognition).
- Filters Multi-cast frames using hash filtering algorithm.
- Discards all frames not addressed to the receiving station when promiscuous mode is disabled.
- Accepts all frames destined to the EMAC if promiscuous mode is enabled.
- Checks incoming frames for transmission errors by way of FCS and verifies octet boundary alignment.
- Discards received transmissions that are less than a minimum length (64 bytes).
- Truncates frames with length greater than maximum frame length when Jumbo frames are not allowed to pass through.
- Optionally forwards pause frames to user application.

B.3 MAC Flow Control Operation

The MAC flow control block is responsible to maintain a proper flow of Ethernet frames through transmit and receive engines. It performs the following main functions:

- Prevents the receive EMAC FIFO congestion by sending pause control frames.
- Prevents the remote device congestion by responding to pause frames and going into idle state for specified number of slot times.

Automatic flow control is only available during the non-PFC (legacy single priority flow control) mode of operation. For PFC mode, user layer is responsible for managing the flow control operation.

B.4 Management Interface

The 10G EMAC core provides a set of signals which can be used to implement the statistics required in IEEE 802.3 basic, mandatory and recommended Management information packages. In addition the MAC core provides signals to generate the applicable objects of the Management Information Database (MIB, MIB II) according to IETF RFC2665.