# TCP Offload Engine

TCP Offload engine implements all the TCP/UDP processing in FPGA logic thus reducing the processing overhead by the system/host. Only the overhead remaining will be the configuration of the logic with right parameters for the socket and interface and DMA transfer of data in both directions. Major features of Vibhatsu TCP offload engine core are
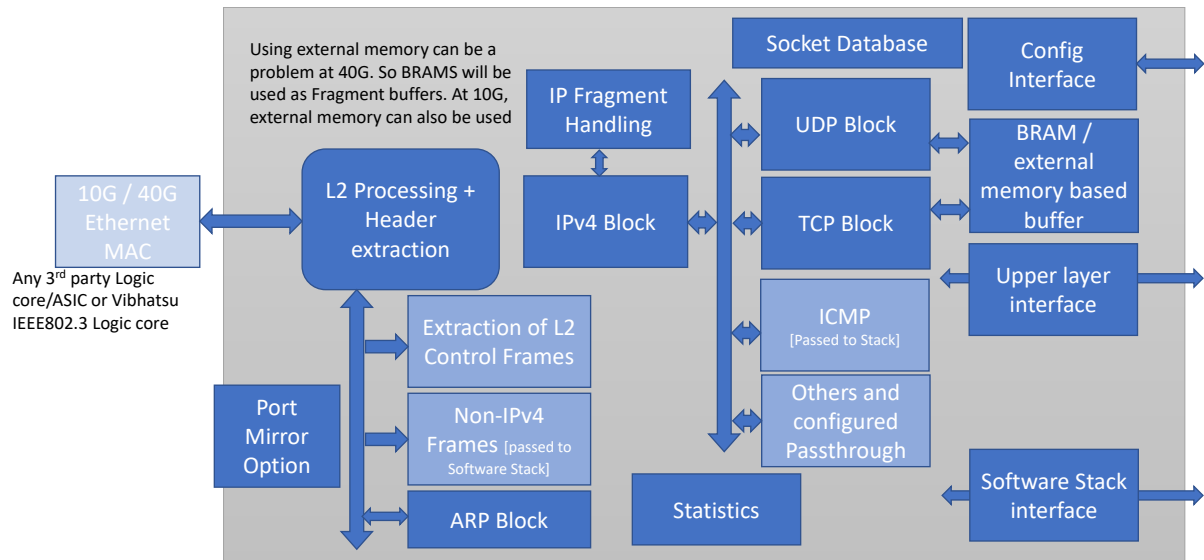
RFCs applocable for the design	
RFC768	User Datagram Protocol [UDP]
RFC791	Internet Protocol ver 4
RFC793	Transmission Control Protocol [TCP]
RFC2018	TCP selective acknowledgement Options
RFC7323	TCP Extensions for High Performance

- Supports 10G with external memory and 40G with internal memory [for 40G, memory bandwidth needed is 5GB or more. It translates to 650MHz or faster DDR @ 72 bit wide bus] [at 40Gbps, max number of frames to process will be 62.5 million per sec per direction. With a roundtrip of 400usec, TCP window size shall be 4MB or bigger]
- Complete implementation of IPv4 block
    - Generation of IP header based on the configured parameters
    - IP header Checksum – compare /insert
    - Marking DSCP [class of service]
    - IP Fragment handling, which is a must for bigger PDUs on UDP [upto 64K byte size PDUs]. [Refer to notes]
    - Moving non configured socket frames to the TCP software stack. This includes non-TCP/UDP [like IGMP or different tunnelled frames or even frames with IP security/encrypted frames]
    - Discarding non-compliant frames
    - Forwards IPv4 frames with extensions to upper layer [only data processing is handled]
    - Statistics as per MIB requirements
- Complete UDP layer

- o Data checksum [Data + pseudo header], discard if error based on config
- o Also implements UDP-Lite protocol [UDP with no data checksum]
- o Frame size upto 64KB
- Complete TCP Layer
  - o Socket initialisation and opening [both direction] based on configuration
  - o Support TCP extensions for High speed [RFC7323] – [bigger window size to address roundtrip delay, bigger time stamp word]
  - o Ack generation, Selective ack option and retransmission whenever necessary
  - o Socket re-opening whenever timeout/reset happens
  - o Upto 32 sockets [for both TCP + UDP. it is limited by the memory space available. For very high speed networks like 40G, 32 sockets is the limit as external memory bandwidth will be a bottleneck]
- Optional features
  - o VLAN insertion and handling – based on config
  - o L2 control frames and Network timing frames can be taken out and sent to the host [part of Access control List]
  - o For smaller socket count [like less than 32] entire logic, database, TCP window buffer and IPv4 fragment handling will be located within FPGA logic and its BRAM [if high speed memory is available, we can use them to have bigger window buffers and fragment buffers]
  - o ARP handling can be optionally included. If not, host has to provide MAC addresses for each of the IP address supported
  - o One test option – like smallest ICMP echo is included for testing purposes
  - o Port mirroring option – entire data can be mirrored to another port, if available on the board for tests
  - o Filtered out frames [frames not processed by the Offload engine] can also be forwarded out through another Ethernet port on the board. [received frames from that port will be forwarded with a slightly lower priority]
  - o Statistics as per MIB requirements for IP, UDP, TCP layers
- Configuration interface – any slow speed interface [another Ethernet or SPI or parallel interface] to configure socket database, addresses and the mini ACL table
- Line side interface – as per IEEE802.3 standard
- System side interface – data path will be parallel bus, while stack interface [for non processed frames can be over Ethernet port [another] or PCIe