10GE TCP Processing

- Current generation NPUs not good match for 10GE TCP processing

- TCP at 10 Gbps characteristics and requirements
  - stateful protocol ➔ efficient RMW
  - large number of connections ➔ scalable architecture
  - jumbled byte stream ➔ intelligent memory system
  - < 10 µs latency requirement ➔ cut-through processing
  - still an evolving protocol ➔ programmability of the implementation
**10GE VLIW TCP Processor**

Terminator Architecture
- Micro-coded VLIW TCP protocol processor
- Up to 1M connections & 4GB buffer memory
- Low latency
  - VeriTest verified 9.7\(\mu\)s user-to-user latency on Linux
  - Through a low latency L2 Ethernet switch!
- High TCP throughput using 1500B Ethernet frames
  - 7+ Gbps through PCI-X 1.0 T110 TOE HBA
  - 14 Gbps in series with IXP2800 10GE Service Blade
- Flat performance profile
  - 7+ Gbps for 1 connection up to 10,000+ connections
- Scalable, ~100% TCP processing efficiency
  - 1 Gbps full-duplex @16.6MHz core clock frequency
  - 7 Gbps full-duplex @125MHz core clock frequency

**10GE VLIW TCP Processor Innovations**

- The TCP protocol is stateful
  - at 10Gbps, 1500 byte packets are 1\(\mu\)sec apart
  - TCB state has poor cache locality
  - wire speed needs to be attainable for 1 connection
  - wire speed needs to be attainable for 1000s of connections

- requires
  - an efficient pipelined pre-fetch of the TCP protocol state
  - single engine that can process 10GE traffic
• The TCP protocol provides a FIFO byte stream abstraction to the end points
  – packets can partially overlap and/or arrive out of order
  – latency requirement rules out store-and-forward

  – Requires
    • specialized memory subsystem to unravel the packet jumble @10GE speeds

• There are stringent end-to-end latency requirements (< 10 μs) in addition to a simultaneous high BW requirement (10Gbps)
  – this requires cut-through processing
  – cut-through processing refers to the packet arriving on one terminal, being processed and then being forwarded out the other terminal without ever being stored in off-chip memory.

• Measured end-to-end latency < 10 μs with L2 switch
  – DMA engine used interrupts but could push number lower by using polling mode
The Need for 10GE VLIW TCP Processor

- Conventional NPU not efficient RMW for stateful protocols
- Byte granularity requires new type of intelligent memory controller
- ~100% TCP processing efficiency
- Shipping today
  - 10GE PCI-X 1.0 HBA with SR & LR optics or CX-4
  - IXP 2800 + Terminator server blade evaluation boards

10GE Service Blade

- 10G and 4x1G
- Up to 512K concurrent sessions
- 4U rack mountable chassis
T110 TOE HBA

- Full TCP/IP Offload
- Supports Full Duplex operation
- PCI-X 1.0a
- iSCSI acceleration
- 10G throughput
- Fiber (Long and Short Reach)
- Diagnostics support
- RDMA (iWARP 0.7)