Low-Power High-Density 10GBASE-T Ethernet Transceiver

Ramin Shirani
Co-Founder
Engineering VP

Ramin Farjadrad
Co-Founder
Chief Architect

Stanford University
August 19, 2011
• Ethernet has been commercially available since year 1980 and today is the dominant networking protocol in the world

• Most common flavor of Ethernet is over twisted pair wire which has evolved from 10Base-T, to 100Base-T, to 1000Base-T, and now to 10GBase-T

• Well over a Billion Ethernet twisted pair ports have shipped

• Standard for Ethernet comes from IEEE LAN / MAN standard committee under 802.3 work group

• Standard naming convention is (data rate) + (base band or broad band) + (media and coding)

  for example: 10GBase-T is 10G Baseband over structured twisted pair cable with LDPC coding
For Server LOM Integration a dual MAC/PHY integrated IC needs to consume close to 10W toward passive cooling.

For dense 48 port switches a low power and very small footprint Quad PHY is important.

It is highly desirable to keep the Phys all in a row and not stagger them:
  - Consistent MDI trace length on all ports
  - Consistent temperature (front to back airflow)

Aquantia’s 25mm x 25mm Quad enables single row:
• High insertion loss 45dB+ @100m & Shannon limit margin <2dB
  • 1GE Cable insertion loss in < 20dB & Shannon limit margin >20dB
• Duplex transmission → Echo power >> Received power
• Self Near End (NEXT) and Far End (FEXT) Crosstalk
• Alien crosstalk from other cables & Environment (EMI)
10GBaseT Technical Requirements

• Low BER (Standard\(<10^{-12}\) & Industry\(<10^{-15}\))

• InterSymbol Interference & Self Noise Cancellation
  • ISI: Cancelled by Rx FFE + Tx THP Filters
  • Echo: Cancelled by Analog & DSP Filters >60dB
  • NEXT: Cancelled by DSP Filters >40dB
  • FEXT: Cancelled by DSP Filters >20dB
  • Radio-Frequency Interference Detection/Cancellation

• Error Correction/Noise Reduction
  • Alien crosstalk/Environment Noise (FEC using LDPC)
  • Circuit Non-linear Distortion & Noise (Hi-Fi Design)
Transceiver Top Level Architecture

- **Control Processor**
  - Sequence Ctrl/monitor/Calibration/Pwr Management

- **MDIO**
  - SRAM

- **Analog**
  - MagJack
  - PLL
  - CDR
  - Xtal
  - Rep
  - DAC
  - A(s)
  - ADC
  - Ecc

- **Digital**
  - FFE
  - FEC
  - LDPC
  - Encode
  - Decode
  - RFI Canceller
  - FEXt Canceller
  - NEXt Canceller
  - Echo Canceller
  - THP Filter
  - FEXt

- **Ethernet PCS**
  - Serdes
  - 10 Tera Operations per Second!

- **Sequence Ctrl/Monitor/Calibration/Pwr Management**
Design Challenges for 10GBase-T

- Low-error-floor forward error correction LDPC
- Efficient implementation of ISI & Self Noise cancellers
- RF interference cancellation (Cell phones/Walkie Talkies) with no performance degradation
- High fidelity data acquisition (ADC&DAC) and clock source circuits at very low power
- On-chip power management to maintain target power over process window
- Extensive firmware to manage power and control modes of operations and interoperability
• **LDPC has an iterative decoder architecture**
  • Provides \(\sim 10\text{dB}\) of coding gain (SNR improvement)
    • Before LDPC: BER \(>10^{-3}\), After LDPC: BER \(<10^{-12}\)
  • Designed for \textbf{Max}=\(\sim 10\) But \textbf{Average}=\(\sim 3\) iterations
    • Over 99.9% of times remaining iteration time is wasted

• **LDPC inherently suffers from Trapping Sets**
  • Limits the coding gain for BER much below \(10^{-12}\)

• **High Complexity: 2048 Equality+384 Parity nodes**
  • Equality: 6 & Parity: 32 message interfaces
  • Message interfaces are bi-directional \(\rightarrow\) 2x8bits
  \(\rightarrow\) Thus \(\rightarrow\) 2048x6x2x8=\(\sim 200\text{K} \text{ wires!}\)
• Less than 0.1% of Codewords need more than 3 iterations!
• Rather than operating LDPC at a clock speed to handle 10 itr, run it half rate to handle 5 itr, while saving the extra 2 itr in a FIFO!
• LDPC at half clock rate allows much lower Vdd to meet timing
• Ideally a logic monitors the FIFO empty depth and adjusts the clock rate and Vdd level according!
• Low message resolution or iterations limits coding gain
  • Message resolution 7-8bit & Iterations 8-10

• “Trapping Set” is a set of codewords that LDPC cannot correct
  • There are ~100 TS codewords thus probability of TS codewords is very low and start to happen at BER <10^{-12}.
  • We have identified & saved on chip TS codewords. Once such codeword is received, it is detected & corrected by a LUT separate from LDPC!
• Transmitting bi-directional on same wire: Cut Wires in half
• Transmitting Serial on same wire: Cut wires by 8 times
  → 16x reduction in # of wires: **200K to 12K wires!**
• Power goes up due to high switching statistics in “serial mode”
• Difference bit transmission eliminates the high switching issue!
ISI Cancellation: 1GE used Rx DFE but 10GE uses Tx THP
- The high latency in 10GE FEC does not allow DFE
  - Transmit data post THP increases 10bits (3bits in 1GE)
  - Noise cancelling filters width increase

Self Noise Cancellation Filters
- Symbol rate: 10GE uses 800MS/s (1.25ns), 1GE 125MS/s (8ns)
  - Filters will be 800/125 times longer
- Cancellation level: 20dB higher in 10GE than 1GE
  - Tap coefficient resolution ~4bits or 50% higher
  - Filter Increase=(10/3*800/125*1.5) > 30x

Must use innovation to minimize Power and Area
- Not all filters taps used all times!
RF Interference Immunity

• **Conventional Approach**
  • Creates frequency notches in Rx filter at the interfering frequencies
  • Filter notches kills signal power too → lower link performance
    • Not very effective on multi-tone or wide-band interferers!
    • Slow response (10-30ms) → Errors happen

• **Aquantia RFI Cancellation**
  • Detects RF interference and directly subtracts it from main signal
  • No filter notches in Rx path → Does not suppress signal power
  • Effective on multi-tone & wideband interferers
  • Almost immediate response → Immune to varying RFI environment
High Fidelity Analog Circuits

• **Low-Jitter Clock Generator**
  - >60dB Echo cancellation requires <1ps RMS jitter!
  - High-Q LC oscillator and regulated clock distribution

• **Low Noise Receive Path**
  - Received signal is over 45dB attenuated
  - PGA+ADC should have >60dB SNR performance

• **High linearity Transmitter/Hybrid (will cover in detail)**
  - Transmit power is over 45dB stronger than received
    → Total Echo cancellation must be >60dB
  - Cancelling filters only cancel the linear portion of Echo
    → Tx DAC/Hybrid must be >60dB linear!
• Single Tx DAC Using Passive Hybrid: PCB traces used as delay lines
  • Low-power as passive hybrid uses no/little power
  • Very sensitive to traces/impedance mismatches
  • Long trace 10inch/channel (40” total!) → Routing issues!
  • Limited risetime control leading to high RF emission power!
    • High capacitance at output pads leads to Return Loss degradations
Two identical DAC design: Uses identical matched DAC to cancel Tx DAC signal

- No long PCB traces $\Rightarrow$ No Mismatch & Routing issue
- Matched DAC burns as much power as main DAC
- DACs distortions will add up $\Rightarrow$ 3dB tighter design spec
- Limited risetime control leading to high RF emission power!
  - High capacitance at output pads leads to Return Loss degradations
Aquantia Tx DAC/Hybrid Architectures

- **Single Tx DAC with Current Mirrors:** Two current mirrors copy DAC signal to line & hybrid
  - Only one DAC w/o off-chip traces → Area & Power saving
  - Ideally 1st Tx symbol distortion is cancelled out
  - No tight matching requirement on DAC or traces
  - Full risetime control before Mirror → Well Controlled RF emission
  - Mirror stage will add additional distortion
Chip Level Power Management

• Process Variations Growing Significantly
  • Leakage is exponentially affected!
    • Over 100x leakage variation
  • Leakage as high as 50% Power

• Leakage Control a Must
  • Back Biasing
    • Performed by on-chip Boost & Negative switching regulators
  • Supply Scaling
    • On-chip miniature linear regulators sprinkled within P&R logic

• Leakage and Speed monitors
  • Large chips need local monitors for better block tuning
• Achieving max channel throughput needs maximum flexibility to fine tune all Analog & DSP parameters
  • Hundreds of controls with Several thousand combinations
  • Using hard-coded state machines creates restrictions

• Firmware Controlled Processor
  • Over 100K lines of code to manage all PHY aspects
  • Managing several modes of operation
    • 10GE, 1GE, 100M, EEE, AutoNeg, Test Modes, interoperability
  • Performance optimization over different channels
    • Optimum Cancellers Adaptation, Analog Tuning, LDPC Tuning
  • Power management across PVT & channels
  • Providing vast debug capability
• The growing use of cloud virtualization is driving the need 10GBase-T as a low cost solution for data centers
• Practical implementation of 10GBase-T Phy requires overcoming many challenges in high fidelity design with outmost attention to low power, small area, and low cost
• Aquantia’s first generation 90nm AQ1002 was introduced in 2009 and now shipping in multiple switching platforms
• Aquantia evaluated analog test chips in 40nm process in early 2009
• 40nm products in the form of Single, Dual, and Quad was first silicon functional and sampled in server and switching platforms in 2010 with production version ramping now