7.3 100Gb/s Ethernet Chipsets in 65nm CMOS Technology

Jhih-Yu Jiang[1], Ping-Chuan Chiang[1], Hao-Wei Hung[1], Chen-Lun Lin[1], Ty Yoon[2], Jri Lee[1]

[1]National Taiwan University, Taipei, Taiwan,
[2]Basas Microelectronics, Beijing, China

This paper presents a complete design of 100GbE chipsets including gearbox TX/RX, LDD and TIA/LA arrays. Figure 7.3.1 shows the architecture, where 10×10Gb/s input data is serialized into 4×25Gb/s bit stream by a 10:4 serializer (i.e., gearbox TX). A 4-element LDD array subsequently drives 4 laser diodes, emitting 850nm light into 4 multimode fibers (MMFs). After traveling over 100m, these optical signals are captured and transformed into electrical domain by means of photo diodes (PDs) and a TIA/LA array. A 4:10 deserializer (gearbox RX) recovers the clock and data, and restores the data sequences into 10×10Gb/s outputs. In applications, gearbox TRX and optical frontends (i.e., LDD and TIA/LA arrays) may be separated by several inches in order to fulfill system-level integration.

The gearbox TX consists of a multi-frequency multi-phase clock generator and two identical sets of 5:2 serializer, each one including five 1:4 DMUXes and four 5:1 MUXes. Modified from [1] with even larger bandwidth, the FFE uses a half-rate structure. A built-in PRBS generator is introduced to facilitate testing, which provides 5 independent 10Gb/s data inputs with length 2^7–1. Gearbox RX design follows the basic structure of [2], which incorporates pre-amplifiers (with gain of 0 to 20dB), mixer-based CDRs [3], 1:10 DMUXes and 4:1 MUXes. The CDR adopts a full-rate linear phase detection structure that has been proven (in silicon) to achieve data rate up to 40Gb/s. In addition to the existing deskew circuit that removes phase offset within 1 bit, we further introduce a bit-alignment circuit in front to fix up to ±7 bits of misalignment due to optical length difference.

The 5:1 MUX circuit is illustrated in Fig. 7.3.2(a), which is realized as a 5-input transmission-gate sampler operated by rail-to-rail data and clocks. Five TSPC FFs with a NOR-gate feedback produce five 20% duty-cycle clocks $CK_{1\sim5}$ for proper sampling. Figure 7.3.2(b) reveals the bit-alignment design, where a 15b shift register governed by 4b control stores data stream. Based on the control signal, 5 consecutive bits out of [–7,..,,0,..,7] can be taken as data output and be set to the subsequent circuit for processing. Since 100GbE only needs fiber channel of 100 to 300m, a 15b phase shifter is sufficient in most cases. To minimize clock phase noise, we realize TX clock generator as a two-stage subharmonically-injection-locking PLL [4] as depicted in Fig. 7.3.2(c). Fixed delays ΔT_1 to ΔT_4 are inserted to ensure that injections always locate at save zone (approximately 187°) over PVT variations.

The optical frontend also involves considerable circuit techniques. Figure 7.3.3 shows the VCSEL driver design, whose core is actually a 2-tap fractional FFE with tunable delay (ΔT) and pre-emphasis factor (α). Here, main cursor (I_0), pre-cursor (I_{-1}), and bias current (I_{DC}) are summed up through M_1~M_5 single-endedly and are fed into the diode. These currents are mirrored from bandgap reference to maintain fixed amounts over PVT variations. Neglecting I_{DC}, the 2-tap FFE's response can be represented as

$$|H(j\omega)| = \sqrt{1+\alpha^2-2\alpha\cos(\omega\Delta T)}$$

$$\text{arc}\,[H(j\omega)] = \tan^{-1}\left[\frac{\alpha\sin(\omega\Delta T)}{1-\alpha\sin(\omega\Delta T)}\right],$$

as illustrated in Fig. 7.3.3. Typical VCSEL requires a compensation of approximately 2dB, i.e., α= 0.25. Considering boosting efficiency and phase concordance, we set ΔT=20ps in this prototype. Built-in 25Gb/s PRBS generator and tunable α and ΔT with a global 12.5GHz PLL are introduced to facilitate testing.

The TIA/LA array employs TIAs, adaptive analog equalizers, and limiting amplifiers, as shown in Fig. 7.3.4(a). TIA must present a well-defined input level for direct dc-coupling from photo diode as well as 50Ω input impedance matching. The proposed TIA is depicted in Fig. 7.3.4(b), where a differential pair $M_{1,2}$ along with local feedback R_F and L_D forms a fixed input level of $V_{DD}-(I_{SS}R_D)/2\approx0.8$V regardless of PVT variations. Note that I_{SS} is properly biased by constant IR circuit, whose on-chip resistor experiences the same variation as R_D does. With R_B and C_B providing unaltered bias point to M_2, signal current I_{in} flows through R_F and currents into voltage as V_{out}. Here, L_D extends the 50Ω matching from 23GHz to 26GHz, and inductive peaking is also used in the output point to further broaden the bandwidth. TIA reveals 40dB gain while consuming 4.8mW of power.

The adaptation of LAs is accomplished by splitting spectrum power and comparing the higher and lower parts, as described in [5]. An on-chip RC filter with corner frequency of 1.5kHz forms offset cancellation loop. The boosting stage is illustrated in Fig. 7.3.4(c), which takes apart the common-source node of $M_{1,2}$ pair and inserts tunable R and C. Each TIA/LA set achieves 72.5dBΩ gain and 21GHz bandwidth with 69mW power consumption. The TIA/LA can actually operate at 40Gb/s, as demonstrated at Fig. 7.3.4(d). VCSEL anodes are powered by a 3.6V supply to arrive at desired extinction ratio (approximately 4 to 5) and thus a symmetric optical data eye. On the other hand, the cathodes of PD array are connected to a 3.6V supply for optimal performance. The VCSEL and PD capacitance are 350fF and 150fF, respectively.

All chips are fabricated in 65nm CMOS technology. The 4-channel Gearbox TRX consumes a total power of 1.84W (TX: 4×200mW, RX: 4×260mW) from a 1.2V supply. The LDD and TIA/LA arrays dissipate 4×99mW and 4×69mW, respectively. The phase noise of 12.5GHz clock in Gearbox TX is plotted in Fig. 7.3.5(a). With injection locking, the rms jitter is only 187fs and the reference spur is –80dBc. The Gearbox TX presents 25Gb/s output data with 1.06ps,rms and 6.67ps,pp jitter [Fig. 7.3.5(b)]. The recovered data at 25Gb/s and 10Gb/s are depicted in Fig. 7.3.5(c), revealing rms jitter of 1.01ps and 2.34ps, respectively. RX jitter tolerance is also shown here. The TX presents a tunable range of 24.6 to 25.74Gb/s, while the RX shows an operation range of 24.94 to 25.19Gb/s. BER < 10^{-12} is ensured through the link. Power penalty for turning on and off adjacent channels is less than 0.63dB [2].

Optical frontend has been tested thoroughly. Figure 7.3.6(a) depicts the LDD's electrical and optical outputs at 25Gb/s. End-to-end test for optical link is also conducted, where the VCSEL emits 1.2mW 850nm light into a 100m MMF for TIA/LA to pick up. With 0.5W/A VCSEL efficiency and 0.47A/W PD responsivity, we obtain the output data as illustrated in Fig. 7.3.6(b). The input referred noise is equal to 4.2μA_{rms}. TIA/LA signal integrity is demonstrated in Fig. 7.3.6(c). It suggests a sensitivity of –6.8dBm optical modulation amplitude (OMA) at 25Gb/s for BER = 10^{-12}, and ±0.12UI opening in the bathtub curve. Figure 7.3.7 shows the die photos. Performance of this work has been summarized and compared with that of other state-of-the-arts recently published.

Acknowledgments:
This work is sponsored in part by Basas Microelectronics and Atilia Technology Inc. The authors thank TSMC university shuttle program for chip fabrication.

References:

[1] H. Wang, C. Lee, A. Lee, and Jri Lee,"A 21-Gb/s 87-mW Transceiver with FFE/DFE/Linear Equalizer in 65-nm CMOS Technology," *IEEE Symp. VLSI Circuits*, pp. 50-51, June 2009.
[2] K. Wu and Jri Lee, "A 2×25Gb/s Deserializer with 2:5 DMUX for 100Gb/s Ethernet Applications," *ISSCC Dig. Tech. Papers*, pp. 374-375, Feb. 2010.
[3] Jri Lee and K. Wu, "A 20Gb/s Full-Rate Linear CDR Circuit with Automatic Frequency Acquisition," *ISSCC Dig. Tech. Papers*, pp. 366-367, Feb. 2009.
[4] Jri Lee, H. Wang, W. Chen, and Y. Lee, "Subharmonically Injection-Locked PLLs for Ultra-Low- Noise Clock Generation," *ISSCC Dig. Tech. Papers*, pp. 92-93, Feb. 2009.
[5] Jri Lee, "A 20-Gb/s Adaptive Equalizer in 0.13-µm CMOS Technology," *ISSCC Dig. Tech. Papers*, pp. 92-93, Feb. 2006.
[6] J. Proesel et al., "25Gb/s 3.6pJ/b and 15Gb/s 1.37pJ/b VCSEL-Based Optical Links in 90nm CMOS," *ISSCC Dig. Tech. Papers*, pp. 418-420, Feb. 2012.
[7] T. Takemoto et al., "A 25-Gb/s 2.2-W Optical Transceiver Using an Analog FE Tolerant to Power Supply Noise and Redundant Data Format Conversion in 65-nm CMOS," *IEEE Symp. VLSI Circuits,* pp. 106-107, June 2012.
[8] G. Ono et al., "A 10:4 MUX and 4:10 DEMUX Gearbox LSI for 100-Gigabit Ethernet Link," *ISSCC Dig. Tech. Papers*, pp. 148-150, Feb. 2011.

 978-1-4673-4516-3/13/$31.00

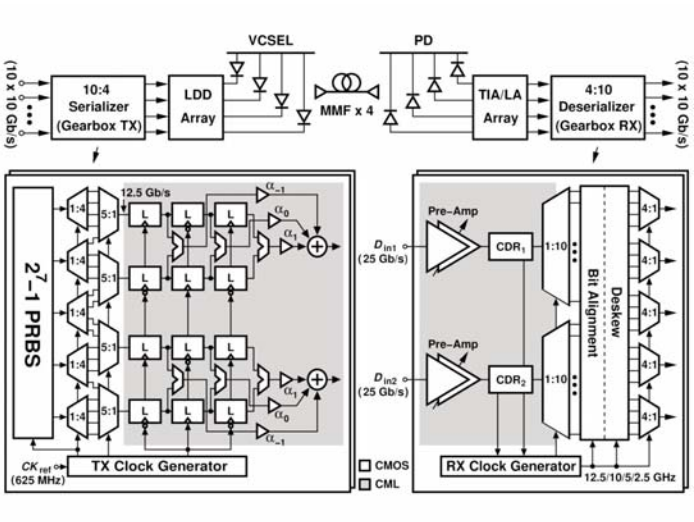


Figure 7.3.1: 100 Gb/s Ethernet architecture.

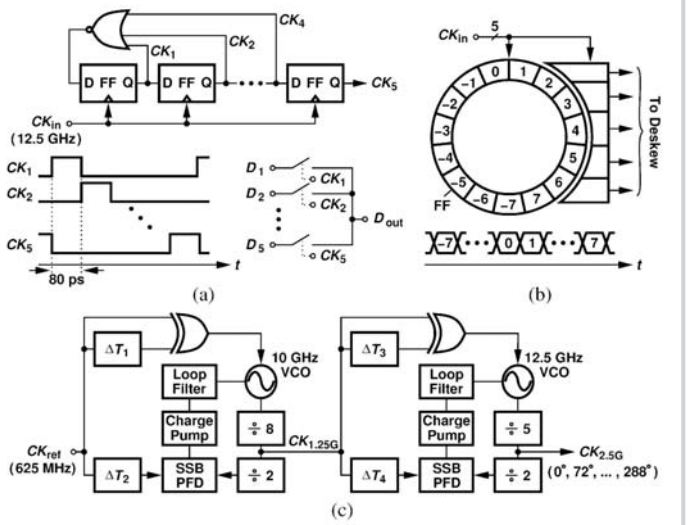


Figure 7.3.2: (a) 5:1 MUX, (b) bit alignment circuit, (c) TX clock generator.

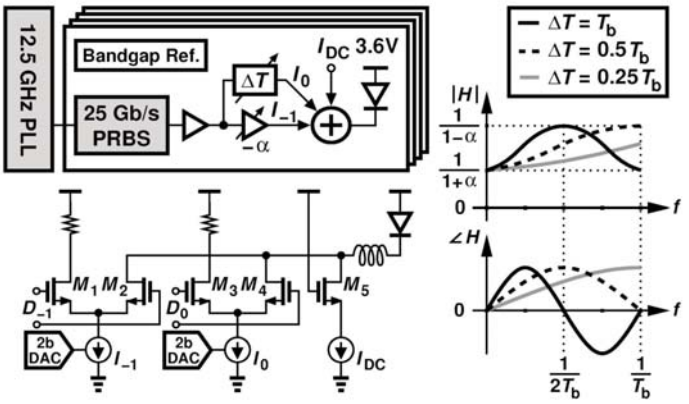


Figure 7.3.3: VCSEL driver array.

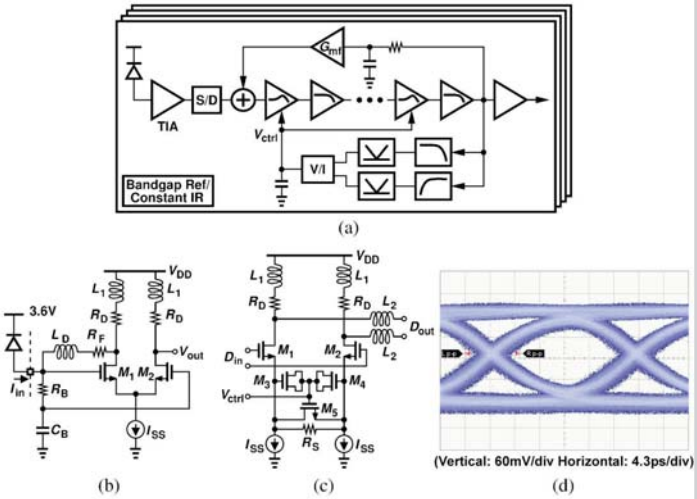


Figure 7.3.4: (a) TIA/LA array, (b) TIA, (c) boosting stage, (d) 40Gb/s output eye @ Vin=47mV.

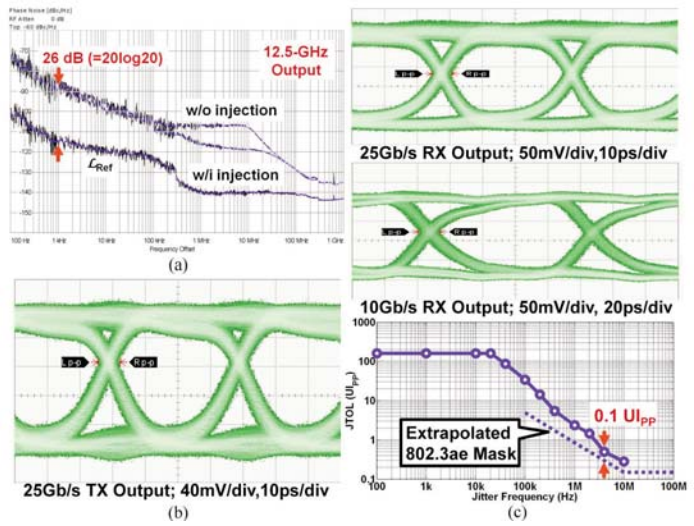


Figure 7.3.5: Gearbox measurements: (a) TX clock phase noise, (b) TX output data, (c) RX recovered data and JTOL.

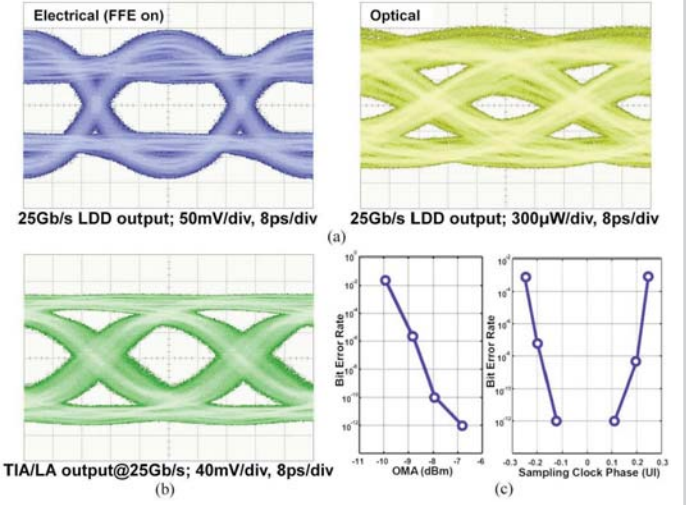
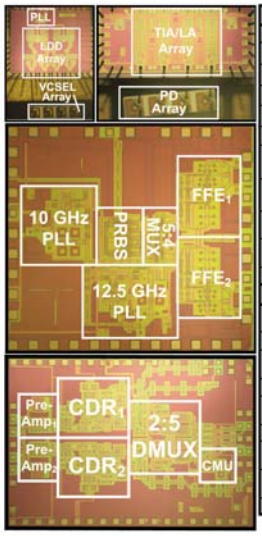


Figure 7.3.6: Optical frontend measurements: (a) LDD electrical and optical outputs, (b) TIA/LA output @ -5.5dBm OMA, (c) TIA/LA sensitivity and bathtub curve @ -5.5dBm OMA.



Gearbox TRX (4 Channels)		
	This Work	[8]
TX Clock PNoise @ 1MHz	−116dBc/Hz	−106dBc/Hz
TX Clock RMS Jitter	187fs (100Hz~1GHz)	429fs (10kHz~100MHz)
Rec. Data Jiiter (25Gb/s)	1.01ps,rms; 6.22ps,pp	N/A
RX Power Penalty	0.63dB	N/A
BER	$<10^{-12}$	$<10^{-12}$
Jitter Tolerance	Exceeds extrapolated 802.3ae mask by $0.1UI_{pp}$	N/A
Tolerable phase error between lanes	±7bit	N/A
Power Consumption	TX: 2 × 400mW RX: 2 × 520mW	1.992W
Chip Area	TX: 1.2 × 1.1 × 2mm² RX: 1.9 × 1.3 × 2mm²	6.3 × 3.7mm²
Technology	65nm CMOS	65nm CMOS

Optical Frontend (4 Channels)			
	This Work	[7]	[6]
Data Rate	4 × 25Gb/s	4 × 25Gb/s	1 × 25Gb/s
Laser Output Power	1.2mW	N/A	0.9mW
TIA/LA Gain	72.5dBΩ	71.2dBΩ	78.3dBΩ*
TIA/LA Sensitivity	−6.8dBm	N/A	−4dBm **
BER	$<10^{-12}$	N/A	$<10^{-12}$
Power Consumption	LDD: 4 × 99mW TIA/LA: 4 × 69mW	LDD+LD: 4 × 208mW TIA/LA: 4 × 59mW	LDD: 1 × 46mW TIA/LA: 1 × 44.4mW
Chip Area	LDD: 1.6 × 1.25mm² TIA/LA: 1.6 × 0.65mm²	Total: 3.6 × 5.3mm²	TX: 0.8 × 0.17mm² *** RX: 0.25 × 0.39mm² ***
Technology	65nm CMOS	65nm CMOS	90nm CMOS

* Simulated Result ** For Data Rate = 22Gb/s, N/A for 25Gb/s *** Core Circuit Area

Figure 7.3.7: Die micrograph and performance summary.

978-1-4673-4516-3/13/$31.00