

A 2×25 -Gb/s Receiver With 2:5 DMUX for 100-Gb/s Ethernet

Ke-Chung Wu and Jri Lee, *Member, IEEE*

Abstract—A 2×25 -Gb/s receiver for 100-Gb Ethernet (100 GbE) has been implemented in 65-nm CMOS technology. A new regulation mechanism is applied to the limiting amplifier to minimize its gain and bandwidth variations. Two low-power full-rate CDRs (with a built-in clock generator) and a high-speed 2:5 DMUX circuit are integrated. Although only two channels are implemented, this receiver provides exactly the same operation as a four-channel one while dealing with independent channels. The prototype achieves bit error rate $< 10^{-12}$ with 20-mV_{pp} input sensitivity, consuming a total power of 510 mW from a 1.2-V supply.

Index Terms—100 GbE, bandgap reference, bit error rate (BER), clock and data recovery (CDR), clock multiplication unit (CMU), demultiplexer (DMUX), deskew circuit, divider, jitter tolerance, limiting amplifier (LA).

I. INTRODUCTION

AS THE ever-growing volume of communication continues to increase, the next generation Ethernet has been moving from 10 Gb/s toward 100 Gb/s. By the beginning of 2010, we have been looking at a few 100-Gb/s standards such as IEEE 802.3ba [1]. Such an ultrahigh bandwidth inspires lots of applications. For example, wireless personal area network (WPAN), network storage, and e-marketing [2]. The 100-Gb/s Ethernet (100 GbE) system includes two data rates (i.e., 40 and 100 Gb/s), supporting full-duplex operation with 64b/66b coding. The standardized transmission media cover both electrical (backplane/copper cable) and optical (single-/multimode fiber, SMF/MMF) channels with various transmission distances. In this paper, we focus on the applications for long-haul communication over SMF. The corresponding specifications have been formulated as 100GBASE-LR4/ER4, stating the channel length to be at least 10 and 40 km, respectively. One of the most critical blocks of this architecture is the physical-medium-dependent (PMD) layer, which partitions the 100-Gb/s optical signal into four subchannels by wavelength division multiplexing (WDM). Operating at 25 Gb/s, this four-lane architecture manifests itself in the improvement of power efficiency and hardware reliability [3]. In addition, the input/output (I/O) data rate on the electrical side remains as 10 Gb/s in order to be compatible with media access controller (MAC). According to the above requirements, the PMD layer should perform the optical/electrical conversion as well as the 4:10/10:4 transformation between 25- and 10-Gb/s data.

Manuscript received March 29, 2010; revised June 21, 2010; accepted July 20, 2010. Date of publication October 14, 2010; date of current version October 22, 2010. This paper was approved by Associate Editor Jafar Savoj.

The authors are with the Electrical Engineering Department, National Taiwan University, 10617 Taipei, Taiwan (e-mail: jrilee@cc.ee.ntu.edu.tw).

Digital Object Identifier 10.1109/JSSC.2010.2074291

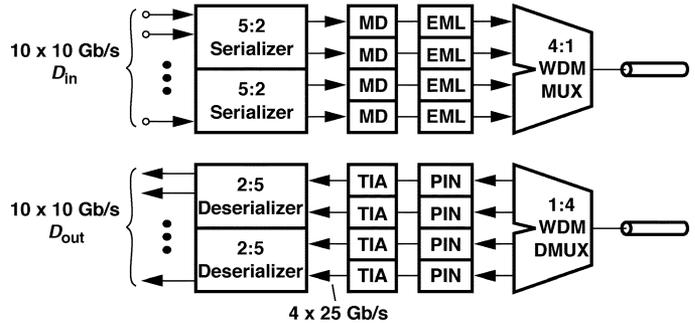


Fig. 1. PMD architecture of 100GBASE-LR4/ER4.

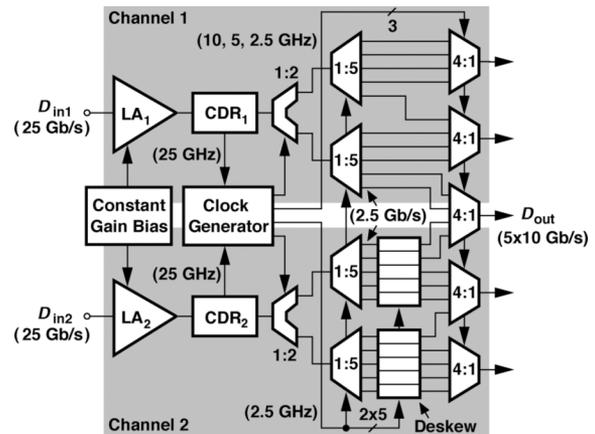


Fig. 2. 2:5 deserializer architecture.

The advanced CMOS technologies continue to provide possible low-power solutions even at such a high speed. Operation beyond 20 Gb/s has been demonstrated in the design of broadband amplifiers [4], [5], clock and data recovery circuits (CDRs) [6], [7], and deserializers [8]–[10]. However, quite a few design issues still remain in the circuit and system levels. First, the gain and bandwidth of limiting amplifiers (LAs) are prone to variation. It makes power optimization very difficult, since the circuits usually need over design to meet the minimum required performance. Meanwhile, CDRs must be able to achieve high speed with low power consumption and small area, which is sometimes contradictory. For example, bang-bang phase detectors (PDs) and subrate architectures are usually adopted to speed up the operation at the cost of higher power and larger area (i.e., complicated routing and a large loop filter). We look into these issues in the following sections.

The 100GBASE-LR4/ER4 PMD structure is illustrated in Fig. 1. Here, ten parallel data inputs are fed into two identical 5:2 serializers to create four outputs, and each of them are 25 Gb/s. After modulator drivers (MDs) and electro-absorption

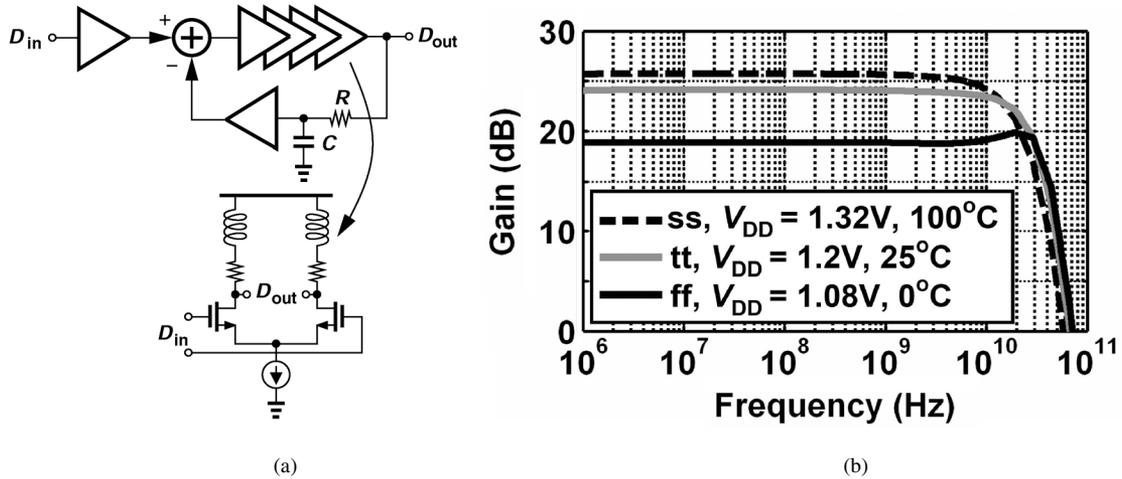


Fig. 3. (a) Conventional limiting amplifier structure and (b) frequency response variations.

modulated lasers (EMLs), the four channels are combined as one optical output by means of WDM. In the receiver part, the input signal is first parallelized as four subrate data streams, and then gets amplified and converted back to the original data format. Note that the two deserializers here are actually operated with the same manner. This structure serializes/deserializes the signals in two steps (i.e., 1:4 and 4:10), providing a compromise for the tradeoff between power and hardware requirements. Among these blocks, the 2:5 deserializer is perhaps the most challenging one. Unlike conventional power-of-2 deserializers (e.g., 1:16 DMUXing in OC-768 [11]), the 2:5 data mapping would suffer from design complexity as well as area and power penalties. Meanwhile, the 2:5 DMUX requires subrate clocks with multiple phases, whose generation and alignment are critical at high frequencies.

In this design, we propose a 2:5 deserializer design fully integrated in 65-nm CMOS technology. The limiting amplifier utilizes a new constant biasing technique to achieve a stable performance over PVT variations. By compensating the device variations, the gain and the output swing are regulated such that the input signals can be amplified to a fixed logic level. The 2:5 DMUX adopts two-step conversion, which substantially reduces the circuit-level design complexity and minimizes the power consumption. It combines building blocks made of CML and digital structures, providing high speed and robustness simultaneously. As will be discussed in the following sections, the complicated function can be realized with simple blocks such as tree-type MUX/DMUX and flipflop-based dividers. This two-step architecture in both data and clock signal flows achieves efficient processing and alleviates the phase alignment issue.

The paper is organized as follows. Section II describes the deserializer architecture. Section III presents the building blocks, revealing design issues and considerations. Transistor-level analysis is included as well. Section IV summarizes the measurement results.

II. ARCHITECTURE

The 2:5 deserializer architecture is shown in Fig. 2. Two channels process the input data independently, presenting an aggregate data rate of 50 Gb/s. Each channel consists of a limiting

amplifier with constant gain biasing, and a full-rate CDR circuit. The two retimed data streams are further demultiplexed into five 10-Gb/s lanes in parallel. The two 25-GHz clocks distilled from the data streams are sent to a clock generator, which creates 2.5, 5, 10, and 12.5-GHz clocks for the subsequent deserializer. Here, we perform an additional 1:2 demuxing right after the CDR to relax the stringent speed requirement. The 1:5 demuxing can therefore be realized in a relaxed way, and finally five 4:1 MUXes are incorporated to produce five 10-Gb/s outputs. A complete 4×25 -Gb/s receiver can easily be implemented by using two identical chipsets proposed here.

The two channels may suffer from significant skew due to channel imbalance. The phase error can be removed by placing a deskew circuit in channel 2, which lines up the 10×2.5 -Gb/s data streams. This adjustment is mandatory because the middle 4:1 MUX has to handle inputs from both channels. Without this realignment, wrong data would be sampled. Note that skews larger than one bit can be removed in system level, e.g., adding first-in first-out (FIFO)¹ buffers, where the maximum tolerable skew range is determined by the amount of registers in the buffer [12], [13]. In this design, we assume an FIFO has already completed the coarse alignment (i.e., offsets greater than 1 bit are removed), and leave only residual skew for our deskew circuit to handle. The channel 2 incorporates deskew circuits between the 1:5 DMUXes and 4:1 MUXes, which automatically pick up proper data phases to suppress the effect of finite skew. This architecture is simple enough to achieve low-power operation.

III. BUILDING BLOCKS

A. Limiting Amplifier

In this design, we need broadband amplifier with at least 25-dB gain and 25-GHz bandwidth so as to provide a large input swing (≈ 500 mV) for the CDR circuits. A conventional approach is shown in Fig. 3(a), where identical gain stages with RC feedback offset cancellation are used. Each stage is realized as a simple differential pair with inductive peaking and a constant tail current. In advanced CMOS technologies, however, device

¹The FIFO design is beyond the scope of this paper and will not be discussed.

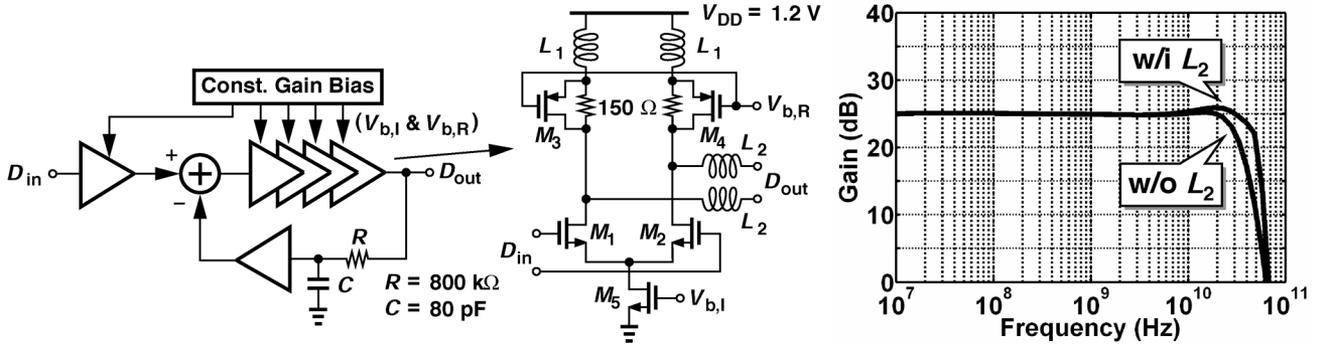


Fig. 4. Proposed limiting amplifier and its response.

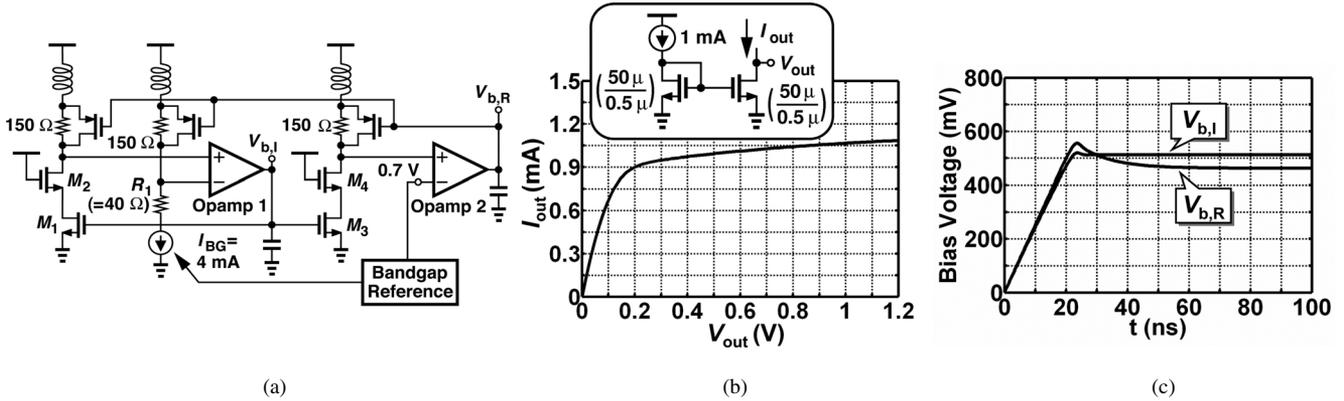


Fig. 5. (a) Bias circuit for $V_{b,R}$ and $V_{b,I}$, (b) simulated output current of a single deck current mirror in 65-nm CMOS, and (c) stability of $V_{b,R}$ and $V_{b,I}$.

characteristics are prone to substantial deviation over PVT variations. For example, if we use 65-nm CMOS to design an LA with 25-dB nominal gain, it can be shown that even with a constant tail current, the overall gain and -3 -dB bandwidth under different conditions still vary by 7 dB and 24 GHz, respectively, for $\pm 10\%$ supply and 100°C temperature variations [Fig. 3(b)]. This issue not only degrades the signal integrity but increases design difficulties in CDR circuits. We thus need more robust biasing circuit to stabilize the LA's performance.

The proposed limiting amplifier is illustrated in Fig. 4. Here, five stages are placed in cascade, and a feedback network is put to suppress offset. The low-pass filter is designed to achieve a low corner frequency of approximately 2.5 kHz, where both R ($= 800\text{ k}\Omega$) and C ($= 80\text{ pF}$) are realized on chip. A constant gain bias circuit dynamically controls the gain and output swing of each stage. The gain stage is realized as a simple differential pair with triple-resonance peaking [4], where the loading ($150\ \Omega$ in parallel with a pMOS resistor) and the tail current ($\approx 4\text{ mA}$) are regulated by the constant-gain bias circuit, i.e., a constant loading resistance ($\approx 125\ \Omega$) and a constant bias current ($\approx 4\text{ mA}$). In other words, the control voltage $V_{b,R}$ and $V_{b,I}$ are regulated by the constant-gain bias circuit. The key point here is to fix the loading ($\triangleq R_{eq}$) and tail current ($\triangleq I_{SS}$) simultaneously. As a result, both the small-signal gain ($= g_{m1,2}R_{eq}$) and large-signal swing ($= I_{SS}R_{eq} \approx 500\text{ mV}$) are fixed over temperature and process variations. The simulated small-signal gain with and without the series peaking inductor L_2 is also de-

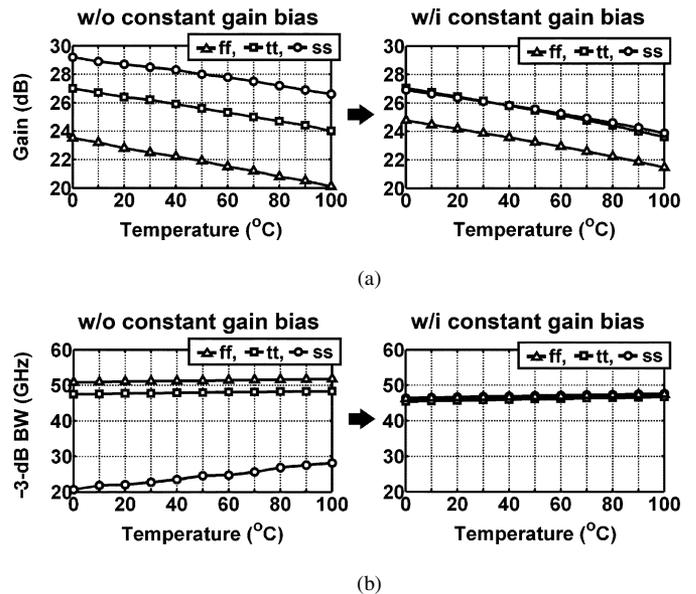


Fig. 6. Improvements of LA variations on (a) gain and (b) -3 -dB bandwidth.

picted here, suggesting that L_2 extends the overall bandwidth by 18 GHz.

The control signals $V_{b,R}$ and $V_{b,I}$ are generated as illustrated in Fig. 5(a). First, a bandgap reference circuit is adopted to create a voltage of 0.7 V ($= 1.2 - 0.5\text{ V}$) and a current (4 mA)

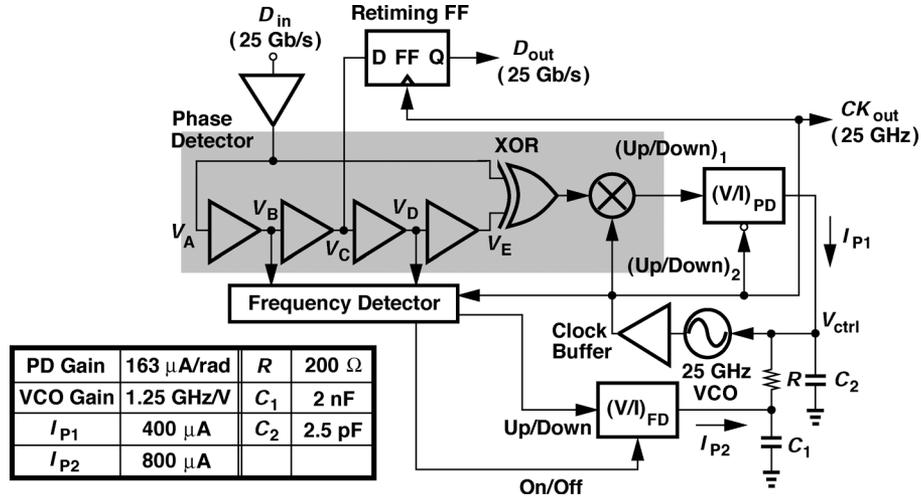


Fig. 7. CDR architecture.

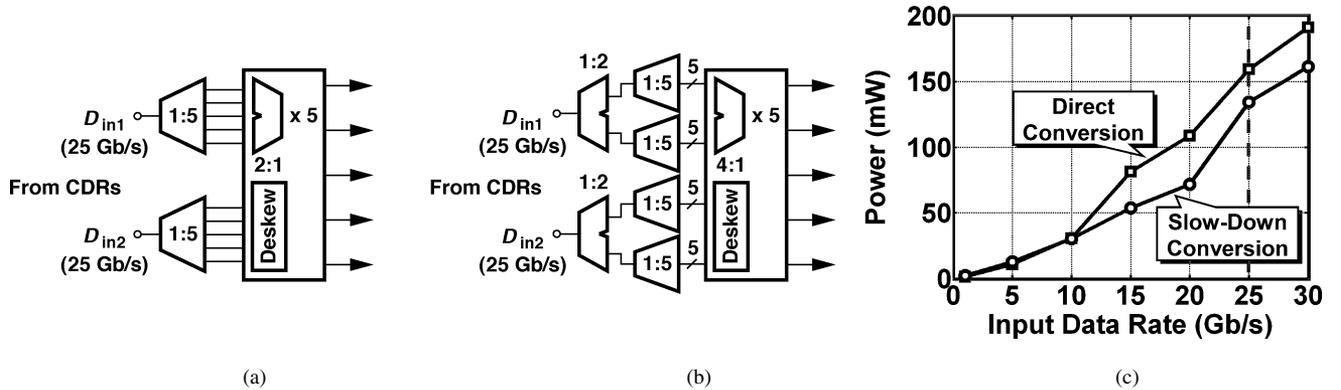


Fig. 8. 2:5 demultiplexing approaches. (a) Direct conversion. (b) Slow-down conversion. (c) Power efficiency comparison.

which are immune from PVT variations. In nanoscale CMOS, a single-deck current source suffers from channel-length modulation when mirroring. As can be shown in Fig. 5(b), even with a channel length of $0.5 \mu\text{m}$, the mirrored current still varies by $14.1\%/V$. Here, we use the loop of R_1 , Opamp 1, and M_1 – M_2 to refine the tail current bias ($V_{b,I}$), which connects to all current sources in the gain stages. It is because R_1 ($= 40 \Omega$) suppresses the excessive V_{DS} of I_{SS} and M_2 mimics the operation of the switching pair. Note that if we were to bias the gain cells with the bandgap current directly, the mirrored current would vary by 8% . The created $V_{b,I}$ also biases M_3 , which together with M_4 and Opamp 2 forms another loop to produce $V_{b,R}$. The equivalent loading resistance is determined by the desired output swing, which is nominally equal to 125Ω ($= 500 \text{ mV}/4 \text{ mA}$). Owing to the existence of the $150\text{-}\Omega$ loading resistors, both feedback loops are stable without any conflict. It is because the Opamp 1 loop will lock unconditionally, and the Opamp 2 loop will follow afterward. At power up, due to the finite loading resistance, which is always between 150Ω (when $V_{b,R} = V_{DD}$) and 90Ω (when $V_{b,R} = 0$) from simulation, the Opamp 1 loop will provide a stable output $V_{b,I}$ first regardless of the transient value of $V_{b,R}$. Once a reasonable $V_{b,I}$ is established, the Opamp 2 loop will be activated afterward. In other words, the two loops are actually

non-interacting. The convergence of these two biasing voltages is shown in Fig. 5(c). Both biases are stabilized to 99% within 55 ns.

Fig. 6 reveals the simulated LA performance under different conditions. The overall gain and bandwidth of the five stages are approximately 25 dB and 47 GHz. Using this compensation biasing, the gain and -3-dB bandwidth variations for different processes and temperatures (0°C – 100°C) are reduced from 9 to 5 dB and from 25 to 2 GHz, respectively. Note that the bandwidth here has been slightly over-designed to ensure a clean eye for the CDRs. Further power reduction can possibly be achieved in future design by optimizing the gain/bandwidth/power tradeoffs. If necessary, supply variation can also be suppressed by a voltage regulator.

B. CDR Circuit

The CDR circuit is shown in Fig. 7, which is modified from [6]. It is basically a 25-Gb/s full-rate CDR circuit employing a mixer-based linear phase detector, achieving high-speed operation by mixing the clock with the data transition pulses. The phase detection here does not involve high-speed pulse generation nor pulse-width comparison. Instead, it creates phase errors in near dc speed and presents an operation range of exactly π . Automatic frequency acquisition loop is incorporated to cover

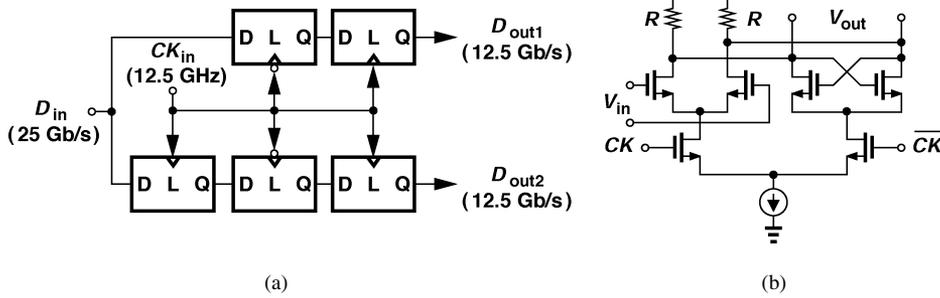


Fig. 9. (a) 1:2 demultiplexer design. (b) CML latch.

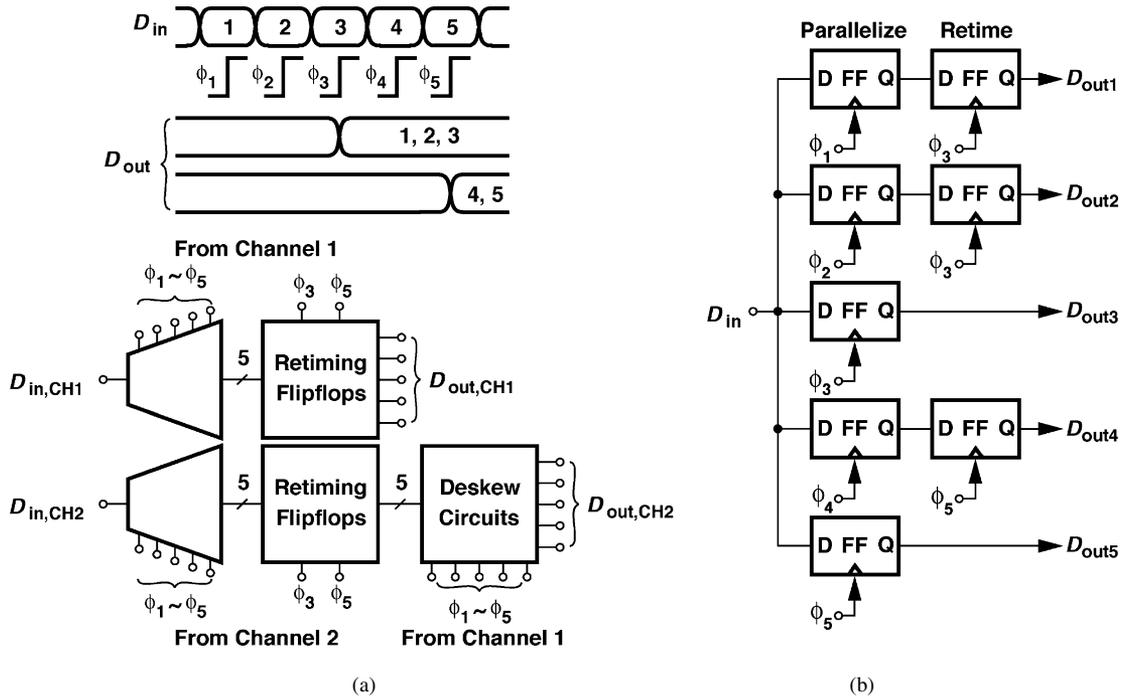


Fig. 10. (a) 1:5 demultiplexing scheme. (b) DMUX with retiming sensing (to ϕ_3 and ϕ_5).

a wide operation range (640 MHz), which utilizes data phases (rather than 90° clock phase) to tell the frequency error. More details can be found in [6]. Note that the CDR has no speed limitation as long as the retiming flipflop and the XOR gate function properly. Simulation suggests the CDR can be designed to operate at 45 Gb/s.

The original clock buffer design consumes more than 40% of total power, so the key modification here is to reduce its power budget while maintaining the same clock swing. With accurate models for active and passive devices, we are capable of using peaking technique more aggressively. For example, the underdamped clock buffer in [6] now provides a peak gain of 6 dB and a -3 -dB bandwidth of 24.4 GHz while consuming only 60% of the original power. Using 65-nm CMOS also helps in power saving, since the overall clock loading has been reduced from 120 to 95 fF. With these optimization, we increase the operation speed by 25% while reducing the power dissipation by 33%.

C. 2:5 Demultiplexer Design

The outputs of the two CDRs are then deserialized into five substrate outputs. We have two possible solutions to do so. Shown in Fig. 8(a) is a straightforward approach, which uses two 1:5 DMUXes to parallelize the two 25-Gb/s data streams into 10×5 -Gb/s lines, and combines every two of them as 5×10 -Gb/s outputs. Such a direct conversion suffers from a few difficulties. First, it is quite stringent to design a 25-Gb/s 1:5 DMUX with reasonable power. In this case one CDR needs to drive eight flipflops (in the first DMUX and divider), which already creates total loading of 17.5 fF and 44.3 fF for output data and clock, respectively, let alone the full rate clock generation and distribution are equivalently challenging. Second, the two sets of lower-speed lines need to be aligned before final combination (2:1 MUXing), and the deskew circuit would consume significant power as well. Finally, the routing of high-speed lines makes the layout even more complicated.

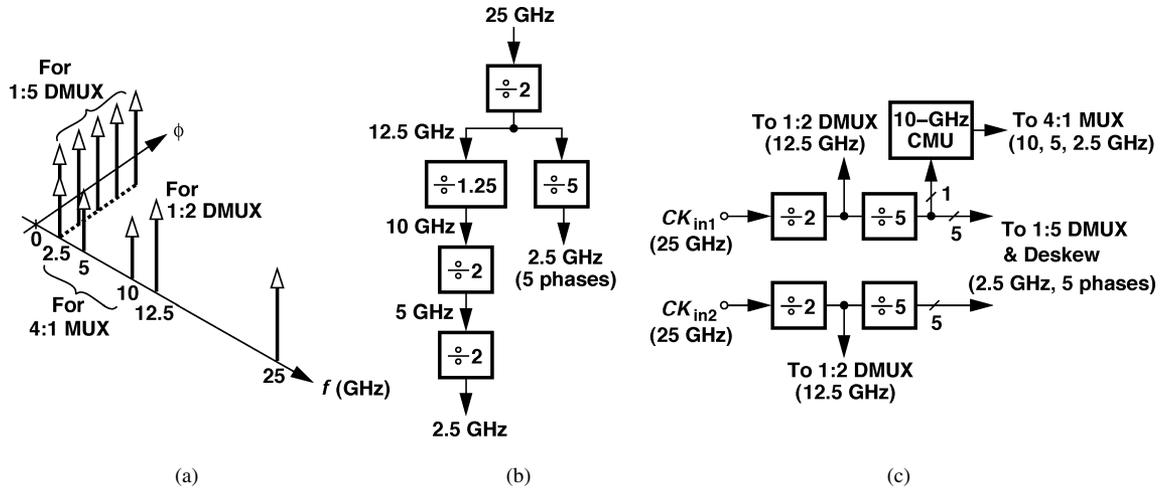


Fig. 13. (a) Required clocks. (b) Illustration of clock generation scheme using fractional dividers. (c) Proposed clock generator architecture.

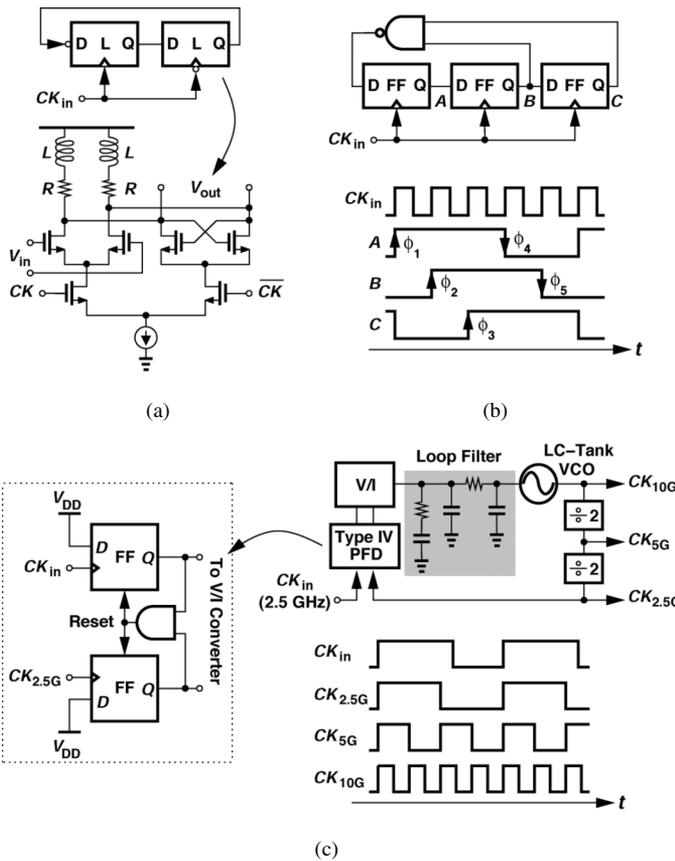


Fig. 14. Implementation of (a) 25-GHz $\div 2$ circuit. (b) five-phase 12.5-GHz $\div 5$ circuit, and (c) 10-GHz CMU.

will be chosen for direct sampling, and vice versa. Fig. 11(b) depicts the output data with and without² the deskew circuit as the input data ($D_{in,CH2}$) coincides with the sampling clock ($\phi_{3,CH1}$). The data eye will be fully destroyed if no deskew circuit is applied.

²Meaning only FF₁ is used.

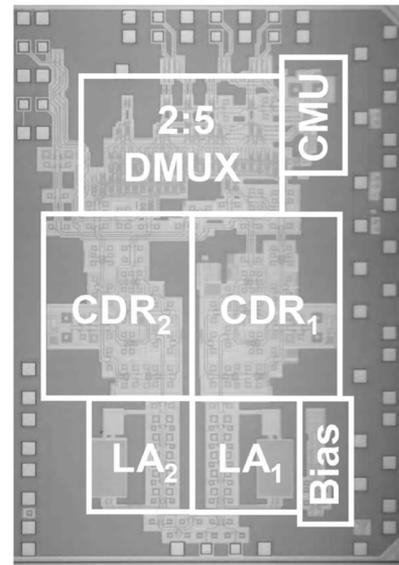


Fig. 15. Chip micrograph.

The 4:1 multiplexer is depicted in Fig. 12. Since the four data inputs have been aligned in the preceding 1:5 DMUX stage, the circuit does not need 2.5-Gb/s shift latches as a conventional MUX does. A 10-Gb/s retimer is placed to clean up the final output data, eliminating possible imbalance caused by data duty cycle error.

D. Clock Generator

The deserializer requires multiple clocks in substrates. As can be illustrated in Fig. 13(a), we need 12.5 GHz for 1:2 DMUXEs, 10, 5, and 2.5 GHz for 4:1 MUXEs, and five-phase 2.5-GHz clocks for 1:5 DMUXEs. One possible way to create a 10-GHz output from a 25-GHz clock is to use fractional dividers [15], [16]. Shown in Fig. 13(b) is one example, where all substrate clocks are generated by frequency dividers. The implementation of the $\div 1.25$ divider, however, would suffer from several design difficulties at high speed. It requires quadrature clocks

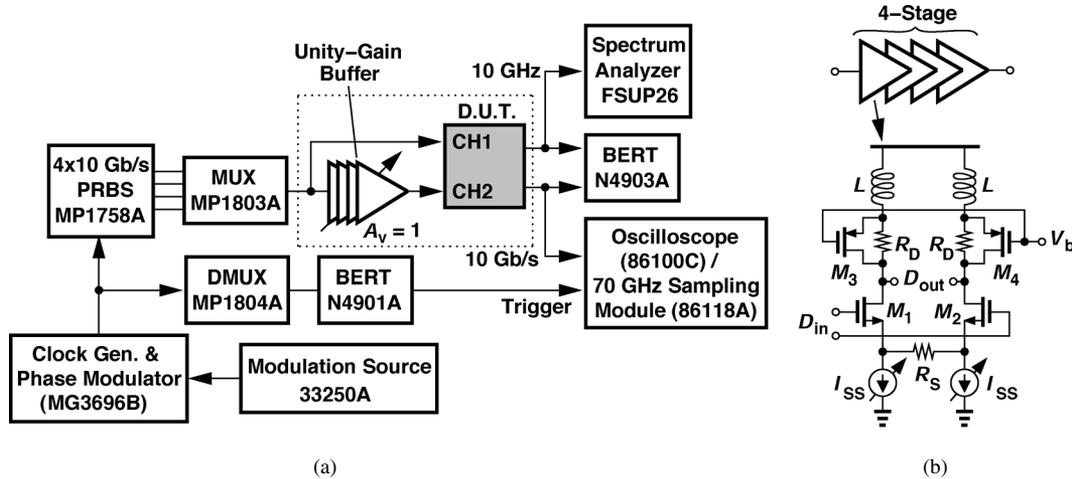


Fig. 16. (a) Testing setup. (b) Design of unity-gain buffer.

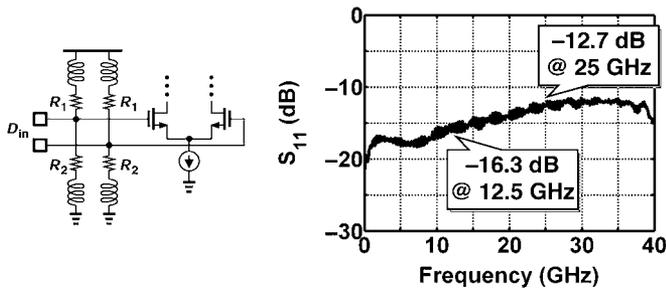


Fig. 17. Input matching network and measured S_{11} of receiver.

of 12.5 GHz to fulfill fractional division, increasing the complexity. The stringent timing on divider phase switching also limits the operation speed to only a few GHz even in 65-nm CMOS technology. As a result, we need a clock multiplication unit (CMU) to create the required substrate clocks.

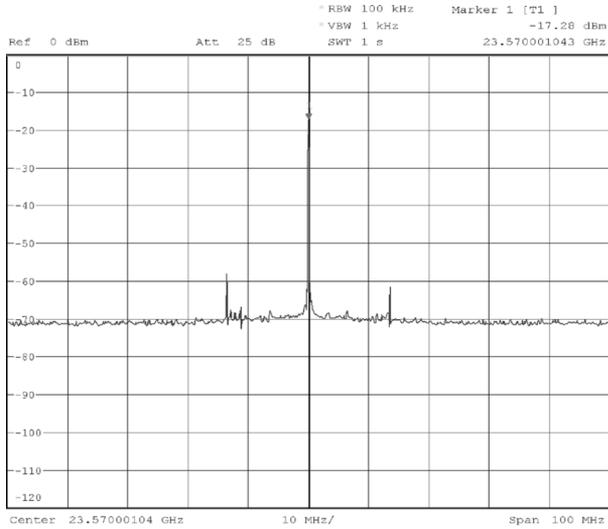
The proposed clock generator is shown in Fig. 13(c). Here, two 25-GHz clocks from CDRs are first divided by 2 and then divided by 5. One 2.5-GHz clock is sent to the 10-GHz CMU to create clocks (with 50% duty cycle) for the 4:1 MUXes. The 10-GHz clock synchronizes all five 10-Gb/s data outputs. The 25-GHz $\div 2$ circuit is realized as a static one, i.e., a feedback flipflop with inductive peaking on the loads [Fig. 14(a)]. The device parameters are identical to the retiming FF in CDR so as to contribute similar CK -to- Q delay. Thus, the clock/phase skew in 1:2 DMUX is substantially alleviated. The five-phase $\div 5$ circuit is implemented as that in [17], where all blocks are made of CML structures including the NAND gate [Fig. 14(b)]. The CMU is shown in Fig. 14(c), which contains a type IV phase frequency detector [18], a corresponding V/I converter, an LC -tank oscillator, and a third-order on-chip loop filter. Here we adopt conventional design to reduce the power and simplify the circuit. Its loop bandwidth is chosen to be 4 MHz, slightly higher than that of the CDR (2 MHz). Such a design allows the CMU to promptly follow the phase variation of CDR and ensures no sampling error in the 4:1 MUX.

IV. EXPERIMENTAL RESULTS

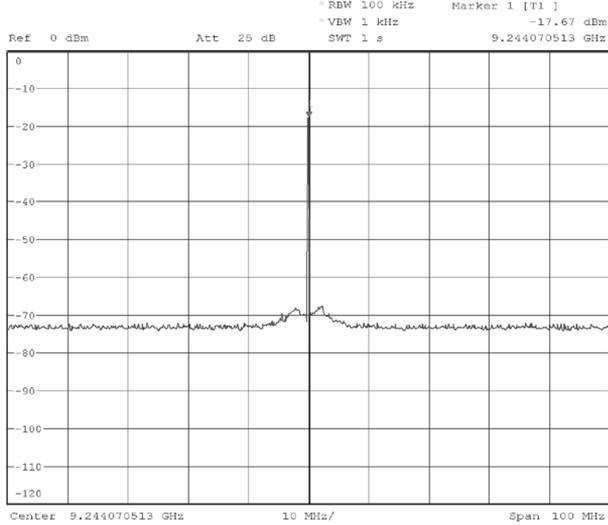
The circuit has been fabricated in 65-nm CMOS technology. Fig. 15 shows the die photograph, which occupies 1.9×1.3 mm² including pads. It consumes a total power of 510 mW for two channels, of which 32×2 mW dissipates in LAs, 99×2 mW in CDRs, 128 mW in the 2:5 DMUX, and 120 mW in the clock generator. A 1.2-V supply voltage is used throughout the chip except the 2:5 DMUX, which requires a 1.4-V supply to accommodate larger data swing. The testing setup is also shown in Fig. 16(a). Due to the lack of two independent data inputs of 25 Gb/s, we place an internal delay to imitate the channel delay. This delay cell must present unity gain to make the two channels experience the same input swing. The unity gain buffer design is shown in Fig. 16(b). It is implemented as a four-stage structure, and each stage contains an inductively peaked differential pair with source degeneration resistor R_S to maintain large linear region. Both tail current (I_{SS}) and loading resistance (R_D with $M_{3,4}$) can be tuned by external control to cope with the gain deviation over PVT variations.³

At 25 Gb/s, the broadband matching at input/output becomes indispensable. The input matching network and the measured S_{11} of the receiver is plotted in Fig. 17, presenting return loss less than -13 dB from dc to 25 GHz. Inductive peaking helps to preserve the 50- Ω matching to much higher frequencies [19]. The CDR reveals an operation range of 640 MHz, across which no performance degradation is observed. Fig. 18(a) shows the CDR recovered clock spectrum (25 GHz) under locked condition, suggesting phase noise of -103 dBc/Hz at 1-MHz offset. The CMU output spectrum (10 GHz) is shown in Fig. 18(b), which demonstrates phase noise of -101 dBc/Hz at 1-MHz offset. Both measurements are conducted with a $2^{31}-1$ pseudorandom binary sequence (PRBS) as the data input. The phase noise plots of the recovered clocks at 25 GHz and 10 GHz are also recorded in Fig. 18(c). It can be shown that the in-band noise of two curves maintain a difference of 8 dB [$= 20 \log(25/10)$], suggesting that all blocks in clock generator contributes negligible noise. The integrated rms jitters from

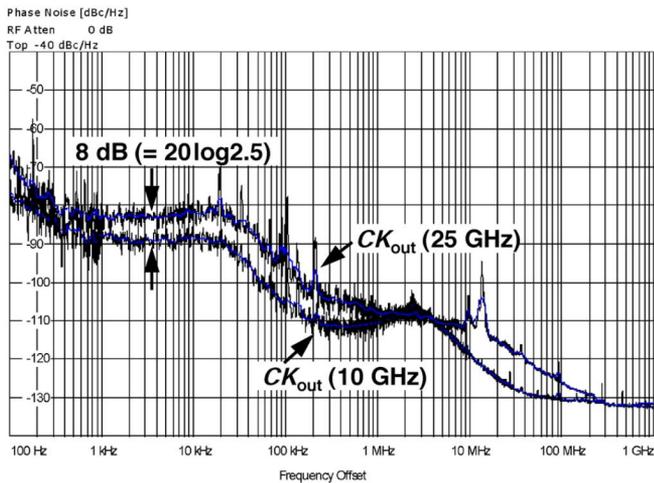
³Due to the limited chip area, the automatic unity-gain regulation is not realized in this prototype.



(a)

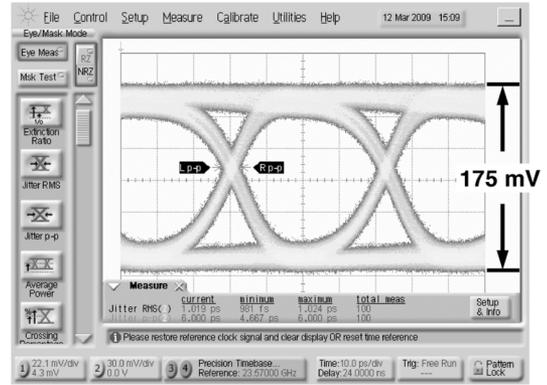


(b)

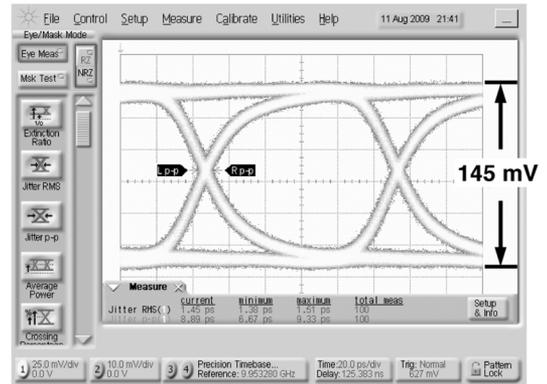


(c)

Fig. 18. Output clock spectrum of (a) 25 GHz and (b) 10 GHz. (c) Phase noise plots. Carrier frequencies deviate from 25 and 10 GHz by a few percents in this design.



(a)



(b)

Fig. 19. Recovered data of (a) 25 Gb/s and (b) 10 Gb/s in response to a $2^{31}-1$ PRBS [horizontal scale: 10 ps/div (left) and 20 ps/div (right), vertical scale: 30 mV/div (left) and 25 mV/div (right)]. The actual input and output data rates are 23.57 and 9.24 Gb/s, respectively.

100 Hz to 1 GHz are 254 and 340 fs, respectively. Note that the CDR’s loop bandwidth here is set to be around 10 MHz. In the integrated receiver, the CDR’s loop bandwidth has been modified to 2 MHz to accommodate with the CMU’s 4-MHz bandwidth.

The 2:5 ratio between input and output data rates leads to difficulties in time-domain and BER measurement as compared with conventional cases. It requires subrate trigger signal for synchronization in order to observe the output data (10 Gb/s). Here, the external equipments (MP1804A and N4901A) serve as a divider chain with total modulus of 20 to create the reference signal for the oscilloscope. Meanwhile, for BER testing, the 5×10 -Gb/s outputs no longer sustain a standard PRBS format. The measurement is conducted by compiling the reference pattern of the error detector (N4903A) in certain sequence so that the output data can be recognized by BER Tester. That is, the BER Tester no longer compares the received data with a standard PRBS, but rather a user-defined pattern. Fig. 19(a) and (b) depict one CDR recovered data (25 Gb/s) and one final output data (10 Gb/s)⁴ in response to a $2^{31}-1$ PRBS, revealing jitter of 1.02 ps, rms/6.00 ps, pp, and 1.45 ps, rms/8.89 ps, pp, respectively. It achieves $BER < 10^{-12}$ for input greater than $20 mV_{pp}$.

⁴The frequencies deviate from the ideal values by 8% and will be corrected in future design.

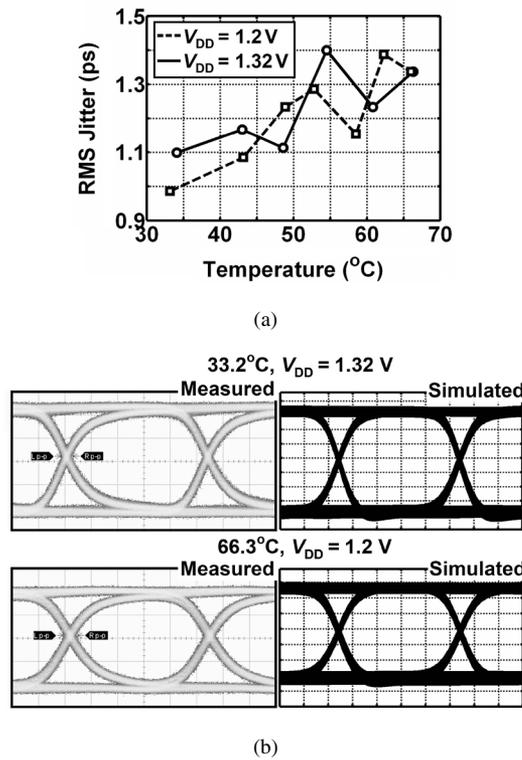


Fig. 20. (a) 10-Gb/s output data jitter for different temperatures and supply voltages and (b) measured and simulated eye diagrams for extreme cases (horizontal scale: 20 ps/div, vertical scale: 40 mV/div).

To check the resistance to PVT variations, we measure the 10-Gb/s output data jitter as a function of temperature and supply voltage. The de-embedded⁵ rms data jitter is depicted in Fig. 20(a), presenting a deviation of 24.7% over 120-mV supply and 33 °C temperature variations. The measured and simulated output data eyes are also shown in Fig. 20(b), revealing negligible difference. Fig. 21(a) shows the jitter tolerance of one 10-Gb/s output in response to a 100-mV_{pp} 2^7-1 PRBS input, which exceeds the extrapolated IEEE 802.3ae mask [21] by at least 0.27 UI_{pp} for all the measurable jitter frequencies with the error threshold set to 10^{-12} . Note that the largest modulation magnitude is 160 UI_{pp}, and the phase variation of our BER tester cannot go beyond 10 MHz. The on-chip interchannel crosstalk is also measured. Fig. 21(b) depicts the channel 1 output BER as a function of input power with channel 2 turned on and off, implying a power penalty of 1.3 dB. Table I summarizes the performance of this work.

V. CONCLUSION

A highly integrated 2×25 -Gb/s receiver with LAs, CDRs, and noninteger DMUXes is demonstrated. Using new biasing technique and optimization in circuit and system levels, the LAs and CDRs present remarkable performance with very low power consumption. A complete design flow of the 2:5 DMUX reveals the benefit of advanced CMOS technology. With these features, we present a promising prototype which kindles the potential of the next generation's 100 GbE.

⁵Noise from oscilloscope is removed [20].

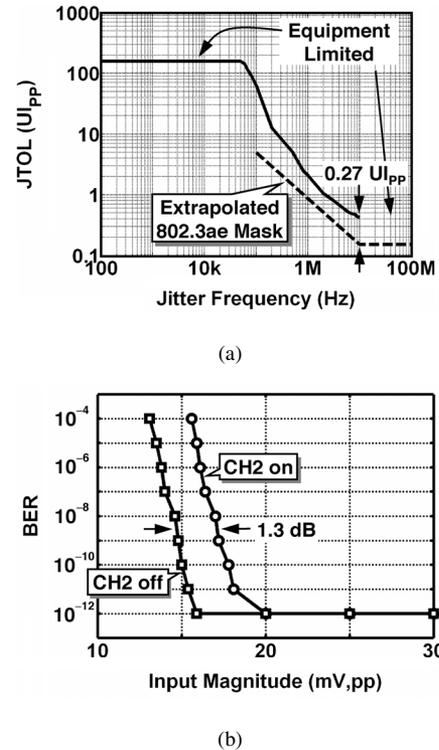


Fig. 21. (a) Jitter tolerance. (b) Power penalty measurement.

TABLE I
DESERIALIZER PERFORMANCE SUMMARY

Input Data Rate	25Gb/s x 2 (23.57Gb/s x 2)
Output Data Rate	10Gb/s x 5 (9.24Gb/s x 5)
Input Return Loss	< -13dB
Rec. Clock Jitter (with $2^{31}-1$ PRBS)	25GHz (23.57GHz): 254fs,rms 10GHz (9.24GHz): 340fs,rms
Rec. Data Jitter (with $2^{31}-1$ PRBS)	25Gb/s (23.57Gb/s): 1.02ps,rms 6.00ps,pp 10Gb/s (9.24Gb/s): 1.45ps,rms 8.89ps,pp
BER (with 2^7-1 PRBS)	< 10^{-12}
Jitter Tolerance	Exceeds extrapolated IEEE 802.3ae mask by 0.27UI _{pp}
Supply	1.2V*
Power Dissipation	Limiting Amps.: 64mW CDRs: 198mW 2:5 DMUX: 128mW Clock Gen.: 120mW Total: 510mW
Chip Area	1.9 x 1.3mm ²
Technology	65nm CMOS

* 1.4V used in 2:5 DMUX.

APPENDIX BANDGAP REFERENCE

The bandgap reference is depicted in Fig. 22(a). Here, the reference voltage V_{ref} ($= 0.7V$) is first generated by the sub-1V

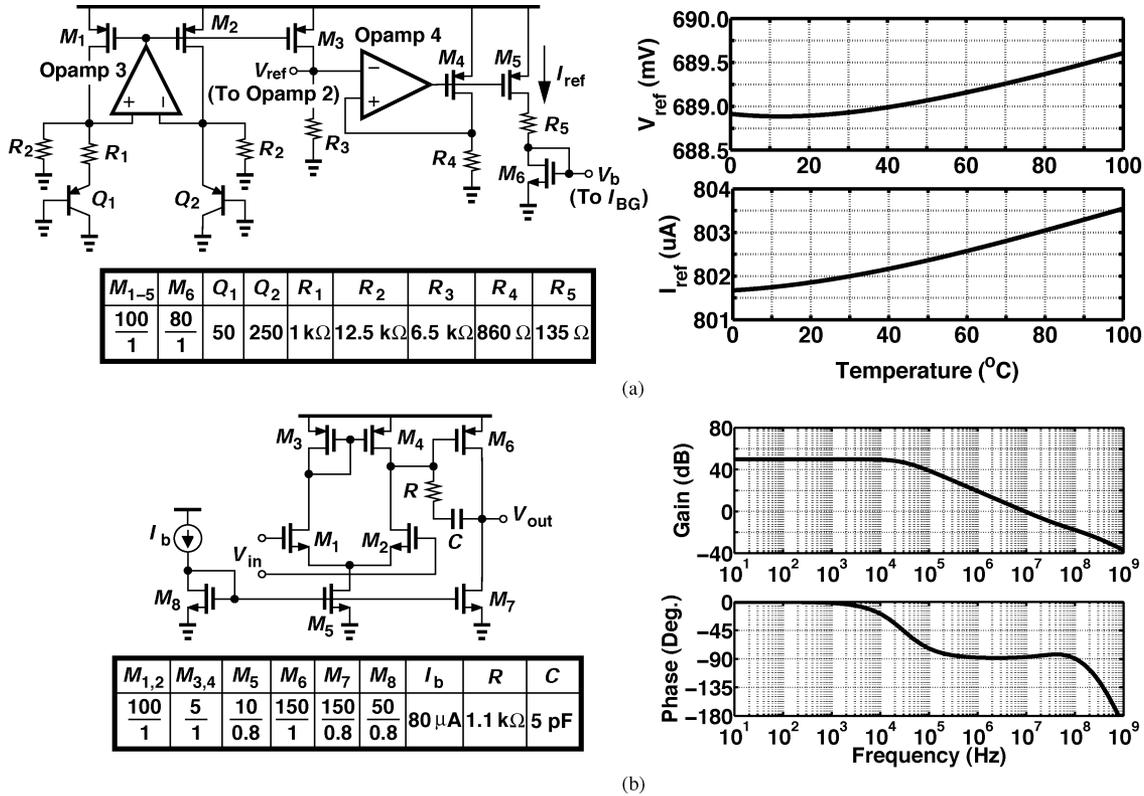


Fig. 22. (a) Bandgap reference design. (b) Its Opamp.

topology. The bandgap current is defined by the subsequent loop of Opamp 4, M_4 , and R_4 , and it is equal to V_{ref}/R_4 . Similar to the bias circuit design, R_5 suppresses the channel-length modulation of M_5 . Following the design in [22], this sub-1V bandgap reference presents V_{ref} and I_{ref} variations of 0.1% and 0.2% from 0 $^{\circ}$ C to 100 $^{\circ}$ C with a power dissipation of 4.2 mW. Fig. 22(a) reveals the variations. The Opamps are implemented as two-stage structures [Fig. 22(b)], which achieve high gain (50 dB) with large output dynamic range.

The loop stability must be handled with care because 1) the two-stage Opamp introduces two internal poles and 2) a third pole exists in each feedback path. To stabilize the loop, we have to push all the nondominant poles away from the origin. First, a compensation capacitor C ($= 5$ pF) and a zero-shifting resistor R ($= 1.1$ k Ω) are placed between the two stages to achieve a large phase margin of 93 $^{\circ}$. Also, to minimize additional phase shift caused by the circuits in the feedback loop, we design the feedback path to have low gain and high bandwidth (i.e., much higher than the unity-gain frequency of Opamp, which is 10 MHz). Simulation shows that all loops maintain overall phase margins greater than 76 $^{\circ}$.

REFERENCES

- [1] 40 Gb/s and 100 Gb/s Ethernet Task Force. [Online]. Available: <http://www.ieee802.org/3/ba/index.html>
- [2] M. Nowell *et al.*, "Overview of requirements and applications for 40 Gigabit and 100 Gigabit ethernet," *Ethernet Alliance*, Aug. 2007.
- [3] C. Cole *et al.*, "100 GbE-optical LAN technologies," *IEEE Commun. Mag.*, vol. 45, pp. 12–19, Dec. 2007.
- [4] S. Galal and B. Razavi, "40-Gb/s amplifier and ESD protection circuit in 0.18- μ m CMOS technology," *IEEE J. Solid-State Circuits*, vol. 39, no. 12, pp. 2389–2396, Dec. 2004.
- [5] J. Kim *et al.*, "Circuit techniques for a 40 Gb/s transmitter in 0.13 μ m CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2005, pp. 150–151.
- [6] J. Lee and K.-C. Wu, "A 20-Gb/s full-rate linear CDR circuit with automatic frequency acquisition," *IEEE J. Solid-State Circuits*, vol. 44, no. 12, pp. 3590–3602, Dec. 2009.
- [7] C. Kromer *et al.*, "A 25-Gb/s CDR in 90-nm CMOS for high-density interconnects," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2921–2929, Dec. 2006.
- [8] K. Kanda *et al.*, "40 Gb/s 4:1 MUX/1:4 DEMUX in 90 nm standard CMOS," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2005, pp. 152–153.
- [9] J.-K. Kim *et al.*, "A fully integrated 0.13- μ m CMOS 40-Gb/s serial link transceiver," *IEEE J. Solid-State Circuits*, vol. 44, no. 5, pp. 1510–1521, May 2009.
- [10] B.-G. Kim *et al.*, "A 20 Gb/s 1:4 DEMUX without inductors and low-power divide-by-2 circuit in 0.13 μ m CMOS technology," *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 541–549, Feb. 2008.
- [11] A. Ong *et al.*, "A 40–43-Gb/s clock and data recovery IC with integrated SFI-5 1:16 demultiplexer in SiGe technology," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2155–2168, Dec. 2003.
- [12] S. Kaeriyama *et al.*, "A 40 Gb/s multi-data-rate CMOS transmitter and receiver chipset with SFI-5 interface for optical transmission systems," *IEEE J. Solid-State Circuits*, vol. 44, no. 12, pp. 3568–3579, Dec. 2009.
- [13] "SerDes Framer Interface Level 5 (SFI-5): Implementation Agreement for 40 Gb/s Interface for Physical Layer Devices," Optical Networking Forum, 2002 [Online]. Available: <http://www.oiforum.com/public/documents/OIF-SFI5-01.0.pdf>
- [14] R. J. Bayrum *et al.*, "A 3 GHz 12-channel time-division multiplexer-demultiplexer chip set," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 1986, pp. 192–193.
- [15] S. Pellerano *et al.*, "A 4.75-GHz fractional frequency divider-by-1.25 with TDC-based all-digital spur calibration in 45-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 44, no. 12, pp. 3422–3433, Dec. 2009.
- [16] C.-W. Lo and H. C. Luong, "A 1.5-V 900-MHz monolithic CMOS fast-switching frequency synthesizer for wireless applications," *IEEE J. Solid-State Circuits*, vol. 37, no. 4, pp. 459–470, Apr. 2002.
- [17] E. Tournier *et al.*, "High-speed dual-modulus prescaler architecture for programmable digital frequency dividers," *IEE Electron. Lett.*, pp. 1433–1434, Nov. 2001.

- [18] B. Razavi, *Design of Analog CMOS Integrated Circuits*. New York: McGraw-Hill, 2001.
- [19] J. Lee, "A 20-Gb/s adaptive equalizer in 0.13- μm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 41, no. 9, pp. 2058–2066, Sep. 2006.
- [20] J. Lee and B. Razavi, "A 40-Gb/s clock and data recovery circuit in 0.18- μm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2181–2190, Dec. 2003.
- [21] *IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements*, IEEE Std 802.3ae.
- [22] H. Banba *et al.*, "A CMOS bandgap reference circuit with sub-1-V operation," *IEEE J. Solid-State Circuits*, vol. 34, no. 5, pp. 670–674, May 1999.



Ke-Chung Wu was born in Taipei, Taiwan, in 1983. He received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 2005. He is currently pursuing the Ph.D. degree in the Graduate Institute of Electrical Engineering, National Taiwan University.

His research interests include phase-locked loops and wireline transceivers for broadband data communication.



Jri Lee (S'03-M'04) received the B.Sc. degree from the National Taiwan University (NTU), Taipei, Taiwan, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Los Angeles (UCLA), both in 2003, all in electrical engineering.

His current research interests include high-speed wireless and wireline transceivers, phase-locked loops, and data converters. After two years of military service (1995–1997), he was with Academia Sinica, Taipei, from 1997 to 1998, and subsequently

Intel Corporation from 2000 to 2002. He has been with NTU since 2004, where he is currently an Associate Professor of electrical engineering.

Prof. Lee serves in the Technical Program Committees of the International Solid-State Circuits Conference (ISSCC), Symposium on VLSI Circuits, and Asian Solid-State Circuits Conference (A-SSCC). He received the Beatrice Winner Award for Editorial Excellence at the 2007 ISSCC, the Takuo Sugano Award for Outstanding Far-East Paper at the 2008 ISSCC, the best technical paper award from the Y. Z. Hsu Memorial Foundation in 2008, the T. Y. Wu Memorial Award from the National Science Council, Taiwan in 2008, the Young Scientist Research Award from Academia Sinica in 2009, and the Outstanding Young Electrical Engineer Award in 2009. He has also received the NTU Outstanding Teaching Award in 2007, 2008, and 2009. He has served as a guest editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS in 2008 and a tutorial lecturer at the 2009 ISSCC.