

A Versatile Low-Jitter PLL in 90-nm CMOS for SerDes Transmitter Clocking

A. L. S. Loke, R. K. Barnes, T. T. Wee, M. M. Oshima¹, C. E. Moore, R. R. Kennedy, J. O. Barnes, R. A. Zimmer, K. L. Arave, H. H. M. Pang, T. E. Cynkar, A. M. Volz, J. R. Pfeister, R. J. Martin, R. H. Miller, D. A. Hood, G. W. Motley, E. J. Rojas, T. M. Walley, and M. J. Gilsdorf

Agilent Technologies, Fort Collins, CO 80525 USA

¹Agilent Technologies, Santa Clara, CA 95051 USA

Abstract

A low-jitter charge-pump PLL is built in 90-nm CMOS for 1–10 Gb/s SerDes transmitter clocking. The PLL employs a programmable dual-path loop filter with integrating path and novel resistorless proportional path that can be independently controlled and accurately modeled for flexible setting of closed-loop bandwidth and peaking. Frequency is synthesized using an area-efficient LC-VCO with helical inductors and inversion-mode nFET varactors for 45% tuning range. The PLL exhibits 0.81 ps rms jitter at 10.0 Gb/s. Technology considerations for improving design manufacturability, tuning range, and jitter performance are addressed.

Introduction

Serial data communication has emerged as the mainstream technology for Gb/s wireline and optical links [1]. This thrust has spawned a myriad of protocols such as Ethernet, Fibre-Channel, PCIe, SAS/SATA, XFI, and SONET that standardize link requirements for the various networking applications. To meet these increasingly demanding protocols cost-effectively, practical *Serializer-Deserializer* (SerDes) I/O channels are frequently required to span multiple data rates, and yet support backward compatibility with legacy link rates and protocols.

This paper presents a transmitter phase-locked loop (PLL) that provides versatile clocking for the outgoing serial data. Closed-loop bandwidth and peaking are adjusted by programming a dual-path loop filter. The transmitter clock is generated by a low-jitter LC-based voltage-controlled oscillator (VCO) that spans a 45% tuning range. This, in conjunction with flexible feedback divider ratios of 10 to 100, enables support of multi-rate protocols including Ethernet (1.25–6.25 Gb/s) and Fibre-Channel (1.0625–8.5 Gb/s). The PLL is integrated into fully embedded 90-nm CMOS SerDes transceiver macros that target networking protocols with data rates ranging from 1.0625 to 10.3 Gb/s. It is fabricated in a standard 90-nm foundry bulk CMOS logic technology that provides 1.0-V core and 1.8-/2.5-/3.3-V I/O FETs with 8-metal copper/low- κ interconnects.

Dual-Path PLL Architecture

The PLL architecture is shown in Fig. 1. It consists of a sequential phase-frequency detector (PFD) feeding the phase error ($\Delta\phi$) between input reference ($REFCLK$) and feedback divider ($DIVCLK$) clocks into a programmable dual-path loop filter. Integrating and proportional paths respectively perform frequency and rapid phase corrections. The integrating path consists of a variable-gain (K_{ip}) charge pump and active integrator driving its VCO input with gain K_{vip} . The operational amplifier feedback forces the charge pump output to $V_{REF2} \approx \frac{1}{2}VDD$ to alleviate charge-pump output resistance limitation. The proportional path consists of a variable-gain (K_{pp}) charge pump driving capacitor C_{pp} with a reset circuit that discharges C_{pp} to V_{REF1} when both $REFCLK$ and $DIVCLK$ are low. This path provides a time-stretched voltage pulse to the VCO with gain K_{vpp} whose average voltage deviation over a $REFCLK$ cycle is equal to $K_{pp}\Delta\phi$.

A differential LC-VCO is employed to achieve low jitter and good supply noise rejection. Each tank is tuned by 159 identical inversion-

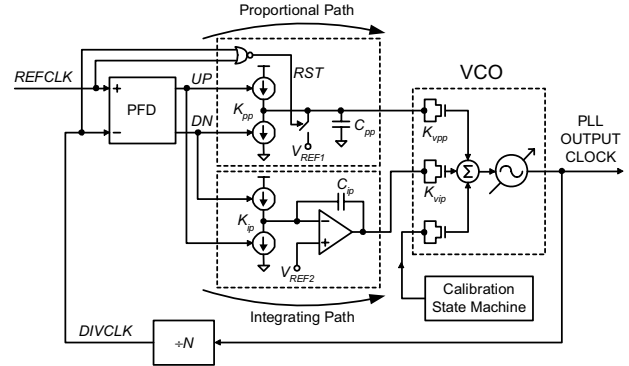


Fig. 1. Charge-pump PLL architecture with dual-path loop filter.

mode nFET varactors [2]: 127 for coarse frequency centering, 12 for analog proportional path tuning, and 20 for analog integrating path tuning. The FETs are configured with gates tied to the tank and shorted source/drain nodes independently controlled. In this way, the VCO conveniently sums the control voltage contributions from all three sets of varactors. Prior to normal PLL operation, a calibration state machine determines the appropriate digital control voltages for coarse-tuning the VCO to its target frequency. Calibration enables support of multiple data rates without compromising loop filter control voltage sensitivity to noise.

The feedback divider has a selectable divider ratio (N) of 10 to 100 in increments of 10 to allow significant flexibility in input $REFCLK$ frequency (f_{REFCLK}). Though not shown, intermediate half- and quarter-VCO-rate frequencies are also derived.

Ignoring higher-order effects, this dual-path PLL closed-loop transfer function can be equivalently modeled as a classical second-order system with damping factor, ζ , and natural frequency, ω_n :

$$H(s) = \frac{\omega_n^2}{z} \cdot \frac{s + z}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (1)$$

where $\zeta = \frac{K_{pp}K_{vpp}}{2\sqrt{NK_{ip}K_{vip}/C_{ip}}}$ and $\omega_n = \sqrt{\frac{K_{ip}K_{vip}/C_{ip}}{N}}$.

In the overdamped case ($\zeta > 1$), it can be shown that of the closed-loop poles and zero (p_1 , p_2 , and z), the pole farthest from the $j\omega$ axis (p_2) sets the closed-loop bandwidth since z effectively cancels p_1 .

$$\omega_{-3dB} \approx -p_2 \approx 2\zeta\omega_n = \frac{K_{pp}K_{vpp}}{N} \quad (2)$$

Eq. (2) shows that the closed-loop bandwidth is primarily a function of the proportional path gain ($K_{pp}K_{vpp}$), independent of integrating path gain ($K_{ip}K_{vip}/C_{ip}$). Furthermore, ζ can be reduced to:

$$\zeta \propto \frac{K_{pp}}{\sqrt{K_{ip}}} \quad (3)$$

to show that closed-loop peaking (inversely related to ζ) decreases for large proportional path gain relative to integrating path gain. This result is intuitively satisfying since in the limit $K_{ip} \rightarrow 0$, the PLL approaches a one-pole system which cannot exhibit any peaking.

Description of PLL Blocks

A. Phase-Frequency Detector (PFD)

Phase comparison is achieved with the PFD shown in Fig. 2. To facilitate VCO calibration and PLL testing, additional control logic is incorporated with minimal latency penalty on normal operation to asynchronously force the *UP* and *DN* outputs to *TEST_UP* and *TEST_DN* respectively when *NORM* is disabled. Using NAND instead of NOR gates in the edge-triggered latches for speed requires a logical inversion that is performed by input NAND gates. These input gates also sharpen the input clock edges to improve noise resilience against slow-rising edges and clock routing parasitic. Incorporating a symmetric NAND in the reset path, wide latch devices, and completely symmetric layout ensures minimal input-referred phase offset error (ϕ_{bs}) between *REFCLK* and *DIVCLK*.

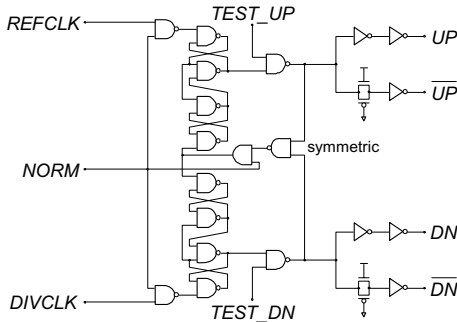


Fig. 2. Schematic of PFD with asynchronous output override.

B. Loop Filter Proportional Path

Phase lead for loop stability compensation is introduced with a resistorless proportional path described in Fig. 3. In dual-path architectures, an obvious approach is to feed a charge-pump current pulse into a parallel *RC* network where the shunt capacitor (*C*) suppresses control voltage bursts that otherwise induce VCO jitter. To save area, we forego the use of a *RC* network. Here, the charge-pump current pulse develops a small voltage shift that is held constant for approximately half the *REFCLK* period and is subsequently discharged to V_{REF1} with a *RST* switch when both *REFCLK* and *DIVCLK* are low. Phase error is not integrated from one phase update to the next since the output is reset to V_{REF1} prior to the next phase comparison. Since the transient pulse width near phase lock is proportional to the *REFCLK* period and corresponding feedback divider ratio (*N*), the voltage pulse area scales with *N* in each phase comparison. The gain therefore self-compensates against differences in divider ratios.

Several charge pump design considerations are important. To achieve good voltage gain for a given charge pump current, the capacitive loading at the charge pump output must be minimized. To reduce charge-pump ϕ_{bs} , we keep the output current source devices in saturation by steering current through a dummy branch when the output node is not charging. A very short delay is introduced when activating this dummy branch to avoid both branches being momentarily cut off during current steering. We also cancel charge injection and clock feedthrough offsets created from *UP* and *DN* switching with half-width switches (not shown) [3] instead of CMOS pass gates in order to mitigate mismatches between nFET and pFET V_T and overlap capacitance. These measures circumvent the need for feedback to match *UP* and *DN* currents. To maximize bidirectional gain linearity across process, supply voltage, and temperature (PVT) variations, we generate V_{REF1} using $\sim V_T$ drop across a large diode-clamped nFET implemented as replicas of the nFET varactors. By tracking V_T , V_{REF1} consistently biases the proportional path output to the center of the varactor depletion-to-inversion transition.

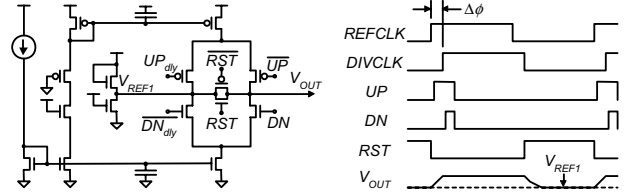


Fig. 3. Schematic of loop filter proportional path and operation.

C. Loop Filter Integrating Path

The integrating path charge-pump gain is selected by a current DAC that generates a binary-weighted fraction of a reference current to be mirrored to the output current source devices. Feedback is employed to match *UP* and *DN* currents in order to minimize reference spurs created by mismatch at low charge-pump currents. As in the proportional path, when the charge pump is inactive, current is steered into a dummy branch in order to maintain the output current source devices in saturation.

The interconnect integrating path capacitor (C_{ip}) is implemented as minimum-spaced interdigitated fingers staggered between consecutive metal levels to maximize interlevel capacitance. Greater area efficiency can be achieved without staggering where fingers are vertically shorted by vias but extra via-to-via capacitance comes at a cost of increased variation in capacitance (hence loop behavior) resulting from typical alignment errors between trench and via lithography in dual-damascene copper processing.

D. Voltage-Controlled Oscillator (VCO)

The PLL output clock is synthesized by an *LC*-VCO shown in Fig. 4. Frequency tuning is achieved with a pair of juxtaposed helical inductors and inversion-mode nFET varactor arrays resonating around *VDD*.

For a given inductance, multi-level helical inductors consume significantly less area as compared to single-level planar spirals due to much tighter mutual magnetic coupling between windings. Although tight coupling exacerbates the proximity effect and degrades coil resistance especially at higher frequencies, higher quality factor (*Q*) inductors are not required in meeting performance targets. Also, despite the use of lower metal levels, capacitive coupling to the substrate is not so severe since the gradual potential drop along the inductor turns leaves the lowest turn exposing only a small fraction of the tank voltage to the substrate.

Inversion-mode or regular nFET varactors were chosen over accumulation-mode devices. First, the gate capacitance vs. control voltage (*C-V*) characteristic is flatter for a control voltage of *GND*, a consequence of normally undesirable poly gate depletion. Since the control voltages for the majority of the varactors (for coarse tuning) are driven to either supply, the *C-V* flatness translates to superior power supply noise immunity and lower jitter. Second, the overall

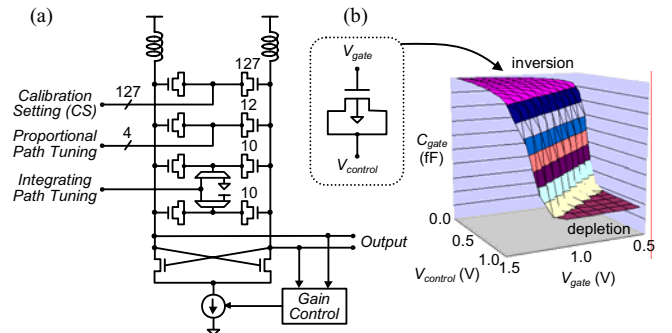


Fig. 4. (a) VCO schematic and (b) *C-V* characteristics of nFET varactor.

tank Q is still limited by the inductor, so higher Q accumulation-mode varactors are unnecessary. Third, nFET C - V modeling is mature and relatively accurate, a critical consideration since the already limited tuning range in LC compared to ring VCOs leaves little to spare for frequency modeling errors. Despite its limitations (such as non-quasistatic modeling), BSIM4 FET models are still more accurate and reliable across PVT as compared to empirical foundry C - V models for accumulation-mode varactors.

Tuning range considerations are obviously important for supporting multi-rate links. Core as opposed to I/O FETs offer the best capacitance contrast between inversion and depletion but raise gate voltage overdrive reliability concerns for large-signal tank oscillations about V_{DD} . As a result, gain-control feedback is incorporated to clamp the oscillation amplitude by limiting the tail node current. Current-limiting the cross-coupled gain FET g_m also improves VCO phase noise by reducing “wasted” current [4].

Better tuning range is attained with longer channel varactors through reducing the relative contribution of overlap capacitance as well as higher substrate capacitance near the source/drain extension due to halo implants that suppress short-channel effects [5]. Moreover, longer (and wider) devices are less sensitive to gate CD variations and provide tighter VCO tuning characteristics. Longer channels do degrade channel resistance and capacitor Q although some Q can be recovered. Extending the length of the source/drain regions reduces mechanical compression in the channel induced by shallow trench isolation [6] and hence minimizes mobility degradation. This also enables landing active area contacts further away from the poly gate edge and nitride gate spacer to additionally improve tuning range; contact coupling to poly gate is an increasingly important parasitic in deep submicron CMOS. To maximize bidirectional tuning in the integrating path following calibration, the tank must be loaded with the average integrating path capacitance during calibration. A pair of analog multiplexers switches half of the varactor inputs to GND and the other half to V_{DD} during calibration, a configuration that overcomes varactor V_T variations across PVT. The feed-forward zero introduced by the multiplexer resistance is negligible.

E. Other Design Considerations

Limited by bump pitch, the PLL floorplan enables unused silicon real estate to be populated with V_{DD} -to- GND capacitance for minimizing jitter induced by supply noise. The I/O nFET with low native doping ($N_A \approx 10^{15} \text{ cm}^{-3}$, $V_T \approx 0 \text{ V}$) is chosen in order to suppress gate tunneling leakage while maximizing stored charge to reduce channel resistance and hence extend filtering bandwidth.

Results and Discussion

A. Closed-Loop Dynamics

Figs. 5 and 6 illustrate a typical example of measured vs. modeled closed-loop bandwidth and peaking. The modeled results demonstrate good agreement to silicon measurements. As predicted by (2) and (3), the closed-loop bandwidth is primarily set by the proportional path gain while peaking is minimized by maintaining a low integrating path relative to proportional path gain. This flexibility enables the PLL closed-loop response to be tailored with good predictability to specifications (bandwidth, $REFCLK$ jitter, spread spectrum modulation, etc.) dictated by the networking standard of interest.

B. Effect of Mismatched Loop Filter Phase Offsets

Architectures employing multiple-path loop filter implementations are prone to higher reference spurs if there is significant mismatch between the input-referred phase offsets of the independent paths. A phase offset in the integrating path charge pump ($\phi_{os,ip}$) will create a steady-state phase offset between the PFD inputs when the PLL is locked. If, however, the proportional path exhibits a different phase

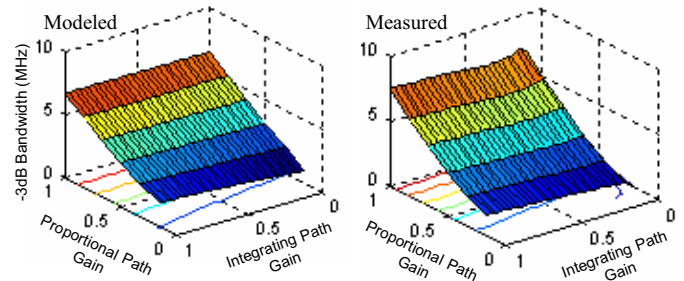


Fig. 5. Modeled vs. measured closed-loop bandwidth as a function of normalized proportional and integrating path gains (1.0 V, 85 °C, $N = 20$).

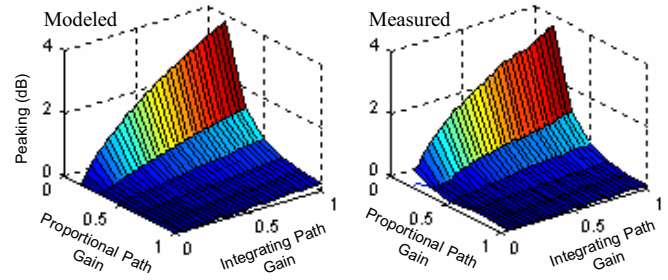


Fig. 6. Modeled vs. measured closed-loop peaking as a function of normalized proportional and integrating path gains (1.0 V, 85 °C, $N = 20$).

offset ($\phi_{os,pp}$), the proportional path will drive its VCO input at every phase update with a voltage pulse proportional to $|\phi_{os,pp} - \phi_{os,ip}|$, thus creating additional VCO jitter at the $REFCLK$ rate when the PLL is locked. Fig. 7 illustrates an example of simulated impact of phase offset mismatch on VCO jitter. The integrating and proportional path charge pumps were designed to mitigate this effect.

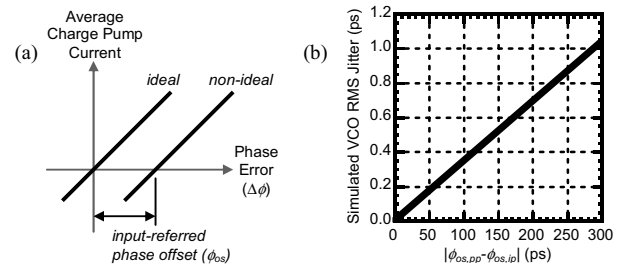


Fig. 7. (a) Input-referred phase offset and (b) simulated impact of phase offset mismatch between proportional and integrating paths on VCO jitter.

C. VCO Performance

The VCO was fabricated with seven inductor variants to target coverage of practical SerDes rates. Fig. 8 shows the VCO coarse tuning characteristics in terms of calibration setting (CS) as a function of

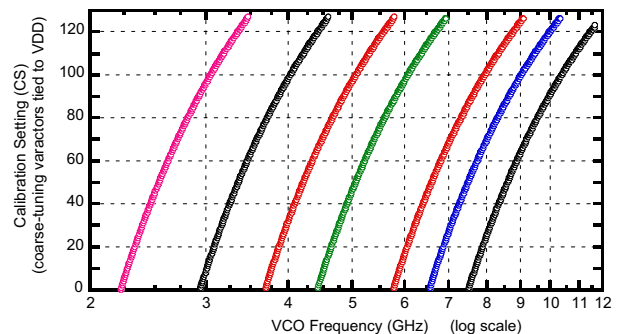


Fig. 8. Measured VCO coarse tuning characteristics for seven inductor variants, each spanning 45% tuning range (1.0 V, 85 °C).

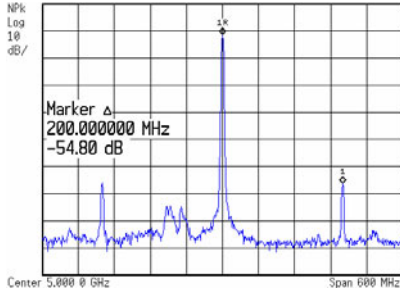


Fig. 9. Measured closed-loop transmitter output spectrum at 10 Gb/s showing reference spurs at $\Delta f=200$ MHz (1.0 V, 85 °C, $N=50$).

VCO frequency ($= N \times f_{REFCLK}$). In this measurement, the analog multiplexers are activated to “midrail” the integrating path capacitance. All inductor variants span a tuning range of 45% which is in good agreement with SPICE prediction of 43%. This match requires FET $C-V$ modeling that accurately reflects the increased substrate doping from halo implants. The transmitter output demonstrates 0.81 ps rms jitter at 10 Gb/s (5 GHz Nyquist rate) with reference spurs at -54.8 dBc (Fig. 9).

The integrating path provides ± 10 varactors of post-calibration tuning to cover VCO frequency sensitivities to VDD and temperature as the PLL operation drifts from the calibration condition. Incorporating excessive integrating path tuning is not prudent as it compromises control voltage susceptibility to noise. We examine the *range* of coarse-tuning CS values at a given frequency across all operating VDD and temperature extremes to verify the PLL robustness in maintaining lock following calibration. The CS range reflects the worst-case capacitance correction needed to maintain frequency lock over worst-case VDD and temperature excursions. As seen in Fig. 10, the CS range does not exceed 10 for all inductor variants at all achievable frequencies. The data further demonstrate that the VCO is more sensitive to temperature than VDD variations, a result attributed to inductor Q limitation [7].

To understand the observed temperature sensitivity, we solve for the series resonance condition of a lossy LC tank, with R_L and R_C parasitics respectively in series with L and C , to obtain:

$$\omega^2 = \frac{1}{LC} \frac{1 - \frac{C}{L} R_L^2}{1 - \frac{C}{L} R_C^2} \equiv \frac{1}{LC} \frac{1 + 1/Q_C^2}{1 + 1/Q_L^2} \quad (4)$$

where the expression in terms of $Q_L = \omega L/R_L$ and $Q_C = 1/\omega C R_C$ is derived via impedance transformation to a parallel tank. With $Q_L \ll Q_C$, we compute dC/dT which is proportional to the measured CS range:

$$\frac{dC}{dT} = -\frac{2R_L C^2}{L} \frac{dR_L}{dT} \approx -\frac{2R_L^2 C^2}{L} TCR \quad (5)$$

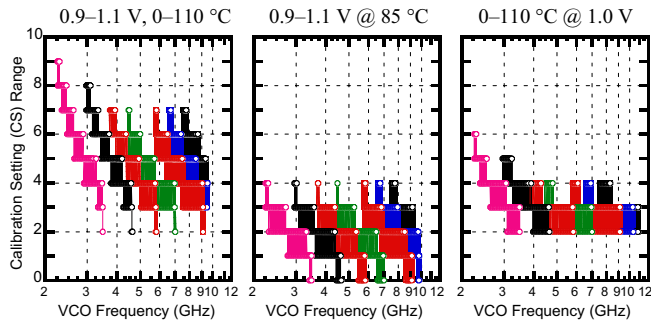


Fig. 10. Measured calibration setting (CS) range across operating supply voltage and temperature (0.9–1.1 V, 0–110 °C) for seven inductor variants.

where $TCR =$ inductor linear temperature coefficient of resistance.

Eq. (5) explains the trend common to all inductor variants in Fig. 10, namely reduced temperature sensitivity at lower C (higher frequency) for a given L . Higher-frequency VCOs with smaller L s did not exhibit worse temperature sensitivity due to a commensurate reduction in R_L made possible by shorting metal turns. The insight in (5) serves as an extra design criterion for optimizing LC-VCOs.

Performance Summary

The PLL measured performance is summarized in Table I. See Fig. 11 for the accompanying die micrograph.

Table I

SerDes Transmitter PLL Performance Summary	
Jitter (1.0 V, 85 °C, $N=50$)	0.81 ps rms @ 10 Gb/s
Reference Spurs (1.0 V, 85 °C, $N=50$)	-54.8 dBc @ 10 Gb/s
Selectable Closed-Loop Bandwidth	0.46–7.5 MHz
Selectable Closed-Loop Peaking	0.0–3.9 dB
Nominal Center Frequency Range	2.9–9.8 GHz
Frequency Tuning Range (1.0 V, 85 °C)	45%
Feedback Divider Ratios	10 to 100 (increments of 10)
Supply Voltage (VDD)	1.0 V (core)
Technology	90-nm CMOS (8M Cu/Low- κ)
Total Power per SerDes Channel	82 mW @ 10 Gb/s
Silicon Area	280 $\mu\text{m} \times 200 \mu\text{m}$

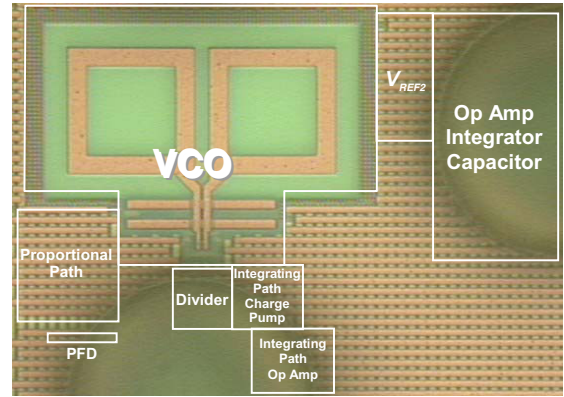


Fig. 11. Die micrograph of transmitter PLL integrated into SerDes macro.

Acknowledgment

The authors thank P. Fisher, R. Hernandez, J. Nguyen, R. Owens, and L. Metz for fruitful discussions. Design support from the SerDes control team, K. Stafford, and T. Abeyta is gratefully acknowledged.

References

- [1] M. Horowitz *et al.*, “High-Speed Electrical Signaling: Overview and Limitations,” *IEEE Micro*, vol. 18, no. 1, pp. 12–24, Jan. 1998.
- [2] P. Andreani and S. Mattisson, “On the Use of MOS Varactors in RF VCO’s,” *IEEE JSSC*, vol. 35, no. 6, pp. 905–910, June 2000.
- [3] B. Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill, pp. 417–423, 2001.
- [4] A. Hajimiri and T. H. Lee, “Design Issues in CMOS Differential LC Oscillators,” *IEEE JSSC*, vol. 34, no. 5, pp. 717–724, May 1999.
- [5] R. Rios *et al.*, “A Three-Transistor Threshold Voltage Model for Halo Processes,” *IEEE IEDM*, pp. 113–116, Dec. 2002.
- [6] R. A. Bianchi *et al.*, “Accurate Modeling of Trench Isolation Induced Mechanical Stress Effect on MOSFET Electrical Performance,” *IEEE IEDM*, pp. 117–120, Dec. 2002.
- [7] T. Tanzawa *et al.*, “A Temperature-Compensated CMOS LC-VCO Enabling the Direct Modulation Architecture in 2.4GHz GFSK Transmitter,” *IEEE CICC*, pp. 273–276, Oct. 2004.