A Semi-Custom Voltage-Island Technique and Its Application to High-Speed Serial Links

Juan-Antonio Carballo[*], Jeffrey L. Burns[*], Seung-Moon Yoo[*], Ivan Vo[*], and V.Robert Norman[†]

[*]IBM Research
11501 Burnet Road
Austin, Texas 78758, USA
+1 (512) 838-8914
(jantonio,smyoo,ivan,jlburns}@us.ibm.com

[†]IBM Microelectronics
3039 Cornwallis Road
Raleigh, NC 27709, USA
+1 (919) 543-6150
robnorm@us.ibm.com

ABSTRACT

Supply-voltage reduction is a known technique for reducing CMOS active power. We propose a semi-custom voltage-island approach based on internal regulation and selective custom design. This approach enables transparent embedding, since no additional external power supply is needed. We apply the approach to high-speed serial links, and we show that high performance is retained through targeted application of custom circuit and logic design. A chip is presented that evaluates the presented approach on a 3000-gate 3.2-Gbps multi-protocol serial-link receiver logic core. When reducing the supply from 1.2V to 0.95V, the chip demonstrates power savings of over 25%.

Categories and Subject Descriptors

B.7.1 [**Hardware**]: Integrated circuits – *types and design styles.*

General Terms

Performance, Design, Experimentation.

Keywords

Low power, communications, serial, links, voltage, island.

1. INTRODUCTION

High-performance intellectual-property (IP) cores are increasingly required to be power-efficient. To achieve high chip-level performance, many of these cores must often be integrated on a common die. Challenging power-performance requirements mandate low-power design techniques that make the most of each technology generation while reducing per-core power consumption.

Unfortunately, time-to-market constraints and design complexity make the option of a full-custom design unattractive. An effective digital power-reduction approach is to reduce the power supply voltage. However, system-level requirements are increasingly requiring various cores to be easily embedded in System-On-a-Chip (SOC) designs, which implies compatibility with standard

technologies and no additional external supplies. These requirements make it difficult to integrate cores that work at lower voltages, since the rest of the SOC will likely include performance-critical components that need standard supply levels.

High-speed communications link cores are critically affected by the abovementioned problem. Many links are often integrated on a single die to increase per-chip bandwidth, but these links are increasingly required to consume low power and to be easily embedded in standard SOCs.

Research on communications links has traditionally focused on producing high-speed low-error-rate analog circuitry. Recent trends require ever lower-power and lower-cost links with increasingly higher proportions of digital circuitry [5][6][7]. Approaches to attack these problems have included advanced circuits and coding techniques [1][4]. Unfortunately, these techniques are becoming increasingly insufficient given the abovementioned power-bandwidth, integration, and time-to-market requirements.

This paper describes a method to use low-supply design techniques inside a core while allowing the transparent integration of such a core in a standard-supply SOC. The method allows for maintaining performance even though the voltage supply is locally reduced. It also allows the use of a standard ASIC methodology; thus a competitive time-to-market can be maintained.

The approach, including new circuits, is illustrated on complex multi-protocol high-speed serial links. Experimental results are shown for a 0.13μm 1.2V CMOS technology.

2. OVERVIEW OF APPROACH

The approach is depicted in Figure 1 for a clock-and-data recovery receiver core. This approach is applicable to many other design problems, including serial link transmitters. It is based on the following techniques:

- **Transparent use of multiple supplies.** Multiple supplies are used to reduce power (see V_{dde} and V_{ddi} in the figure), while integrated low-voltage regulation is applied so the core can be integrated on a chip as if it were a single-supply core. Power supply routing is simplified as these analog and digital blocks are treated as large single-supply cells.

- **Selective custom design.** Custom logic design is selectively applied in such a way that the performance of low-voltage

portions is maintained and the use of standard ASIC methodologies is enabled.
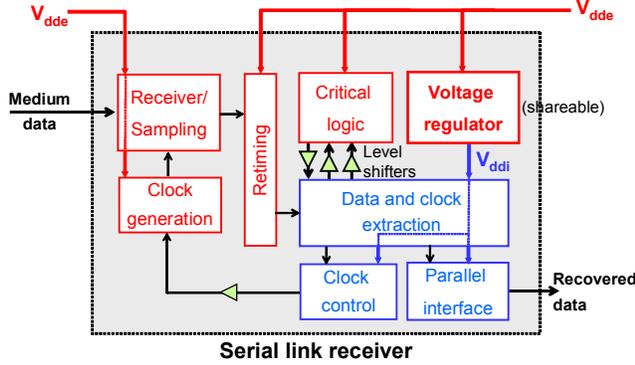
Figure 1   Transparent-supply serial communications link receiver.

The approach is described in more detail in the next subsections.

2.1   Energy efficiency

Power consumption is significantly reduced by powering a large portion of the digital logic at a low supply voltage, while the analog portion runs at a higher supply level.

As shown in Figure 1, most of the communications link logic, including data decision and extraction, clock recovery, sampling clock control, and serial-to-parallel conversion are powered from a down-converted power supply. Analog or performance-critical functions, such as over-sampling, clock generation, retiming and other critical logic are powered at standard supply levels.

2.2   Ease of integration

To make this technique transparent to the SOC in which the core is embedded, an on-chip regulator is embedded in the core to down-convert the high input voltage supply to the low internal supply. (Note that more internal power supplies are possible.) The number of pins and the power supply distribution in the SOC are thereby unaffected. For on-chip regulated digital logic, the power consumption, denoted by P, can be approximated as

$$P = V_{dde} \cdot I_{dde} \cong V_{dde} \cdot (I_{reg} + K \cdot f \cdot V_{ddi}) \ (1)$$

where V_{dde} and I_{dde} are the external power supply and consumed current, respectively; V_{ddi} is the down-converted voltage level; I_{reg} is the current consumed by the regulator; f is the average frequency at which the logic runs; and K is proportional to the average switching capacitance. Thus significant power savings are possible as long as efficient on-chip voltage regulators are used.

Linear regulators are generally the only realistic option for on-chip regulation [2]. Using these regulators, and assuming high current efficiency, the power savings are approximately linearly proportional to the reduction in voltage:

$$P_{savings} \propto (1 - \frac{V_{ddi}}{V_{dde}}) \ (2)$$

This relationship can be explained by the need to drop the voltage (V_{dde}-V_{ddi}) through the linear regulator's pass transistor device.

Note that using on-chip regulation also results in tighter supply corners. As a result, the effect of "fast-chip" corners on power consumption is mitigated, thanks to a reduction in the number of timing races[1].

2.3   Performance

For a high-performance core, it is critical to maintain the core performance when reducing its power consumption. Specifically, commercial high-speed links target a set of markets where bandwidth performance is often fixed by industry standards. In order to recover any performance loss due to reducing the supply voltage, the following techniques are used:

- The logic is first re-optimized using synthesis- and sizing-based optimization.

- Custom design techniques are then applied to difficult paths in order to maintain performance while retaining the power savings.

- For certain logic portions including the most critical paths, a higher-voltage supply is used (shown in Figure 1 to be the same as the analog portion's level for simplicity), and a set of scannable level-shifter latching circuits convert logic levels across the supply-voltage domains. The position of the border between the two voltage domains can be optimized to minimize the number of interface points and thus the level-shifting overhead.

2.4   Area

The area impact of this approach is small and decreases with the number of functions (e.g., communications links) in a core. The area overhead of the voltage regulator is minimized by sharing one regulator across multiple functions. While the area of the output section of the regulator grows with the load it must drive, it typically has other circuit elements, such as a voltage reference generator, whose sizes do not vary significantly with the load.

The area overhead of the voltage regulator can be under 5% when it is amortized over a four-link core. This technique is very applicable to serial links, since communications SOCs often contain tens to hundreds of links integrated on a single die.

---

[1] Adaptive regulation may be used to help reduce process and temperature effects further.

3. SELECTIVE CUSTOM DESIGN

The following selective custom design techniques are applied so the core maintains its performance at low supplies while enabling the use of ASIC methodologies:

- **Critical path length reduction.** Low-voltage critical paths are shortened through application of custom microprocessor-style design techniques. Specifically, logic functions are aggressively integrated into edge-triggered latches resulting in improved delays, set-up times, and hold times. Manual logic optimization and sizing are also applied.

- **Multi-threshold logic.** Custom cells using carefully-placed low-threshold transistors are selectively used to optimize the power-area trade-off while maintaining high-supply-like performance.

These selective custom design techniques are illustrated in Figure 2.
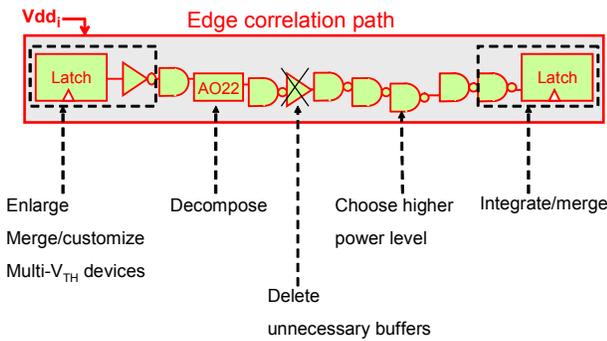
Figure 2   **Example custom techniques.**

Figure 3 shows an edge-triggered mixed-V_{th} integrated NOR+latch that works at under-1V voltage supplies.
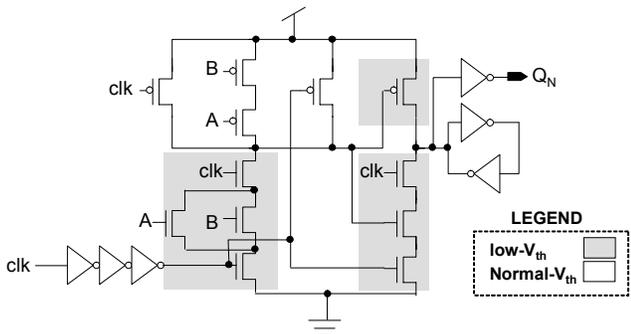
Figure 3   **Integrated NOR-function latch.** *A* and *B* are logic inputs, and *clk* is the clock.

Integrating logic and storage in the latch complicates its delay path by adding complexity to critical transistor stacks. To address this issue and optimize the area-power-delay trade-off, complexity is concentrated on NMOS stacks, and low-threshold devices are used selectively in critical devices, particularly certain PMOS devices. Specifically, the clock-to-output critical path includes the first-stage NMOS stack and the second-stage PMOS device.

Integrating logic functions into latches dramatically reduces delays in the critical path at low supplies by literally cutting down the number of logic stages. The required performance is thereby achieved at much lower power consumption, with little area penalty. Since the latch is edge-triggered, no change in the clocking methodology is needed.

Since scanning the voltage-domain boundaries is necessary, another type of custom cell was designed (see Figure 4), a fully-scannable LSSD level-shifting latch that works down to 0.8V input.
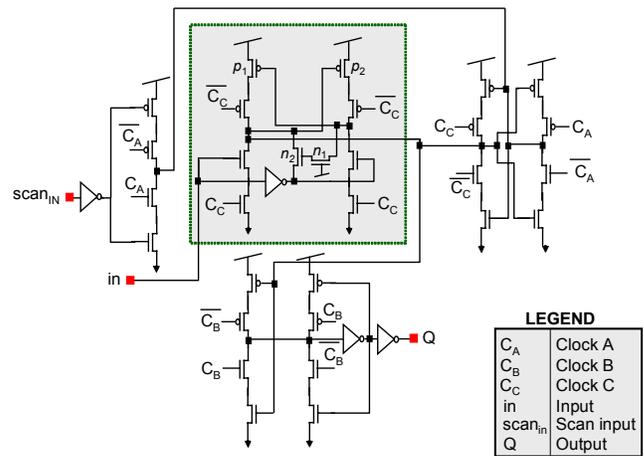
Figure 4   **Scannable LSSD level-shifting latch.**

This circuit consumes 50% less power than a scannable latch connected to a standard level shifter with no loss in performance. The area impact is insignificant, when compared with the original latch. Clocks C_B and C_C are used to transfer data in non-scanning mode. (The input-stage PMOS devices connected to C_C are optional.) The latch is fully scannable using clocks C_A and C_B, and it is V_{dde}-powered except for the input inverter.

Level shifting is done at the input stage in a differential fashion, thereby guaranteeing a full high-level output voltage. Transistors n_1 and n_2 are inserted to reduce short-circuit current in the differential stage. Before the input goes low, n_2's gate is precharged. When the drain of n_2 goes high, its gate is bootstrapped. The source of n_2 is charged to a full supply until n_2 is turned off. Device p_2 is thereby turned off more strongly, thus reducing short circuit current (Note that other embedded level-shifting techniques are also possible.)

Figure 5 illustrates the improvement in instantaneous current consumption for normal operation achieved by using new latches when compared with the conventional approach (conventional cross-coupled shifters connected to standard LSSD latches).
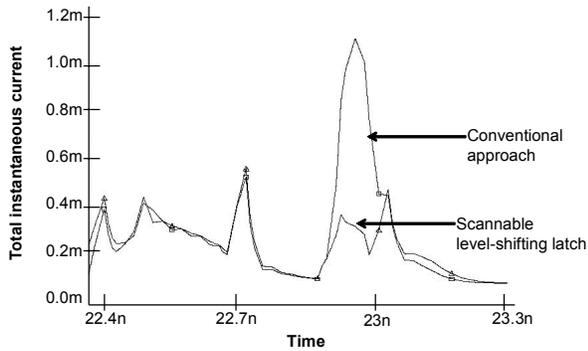
Figure 5   Instantaneous current comparison for scannable LSSD level-shifting latch.

Since the custom techniques described above can be packaged as conventional static standard cells, these techniques can be readily applied in a standard ASIC methodology environment. The number of custom cells needed for a typical production link in our experience is only 10-15 cells.

4. SUPPLY CHOICE

The abovementioned custom techniques enable the use of synthesis-based ASIC methodologies at very low supply voltages. However, a key issue to be addressed is the exact level of supply at which to operate.

Figure 6 shows the delay-power trade-off in a key critical path when using the selective custom techniques above.
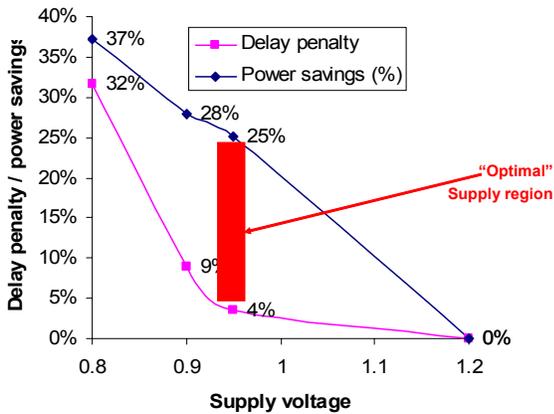
Figure 6   Critical path-based supply choice.

This simulated data suggests an optimal supply of 0.95V (for a 1.2V nominal technology), since at this level significant power savings are possible at practically no performance penalty. (Further reduction in voltage results in a large degradation in delay.)

The majority of the logic has some slack, however. Therefore, most of the logic does not require custom optimization and results

in even higher power savings. (Note that the remaining 4% delay penalty exceeds the precision of pre-layout-optimization simulation. It is addressed by block-level post-layout optimization.)

5. SUB-1V REGULATION

For this technique to be applied in production cores, including links, sub-1V CMOS on-chip regulation with outstanding temperature and transient response is needed. However, it is hard to get a low reference voltage with accurate temperature compensation using standard CMOS regulators.

Figure 7 shows the regulator approach used. A high-gain amplifier is employed. The NMOS driver is a zero-threshold device that allows small differences between input and output voltages and mitigates the supply noise produced at the V_{dde} pin.
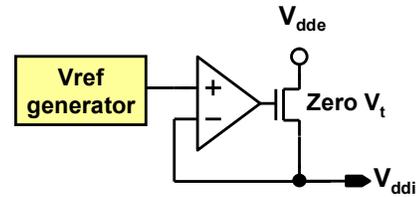
Figure 7   Linear regulator approach.

Figure 8 describes the approach used to generate a sub-1V reference voltage, based on [1]. A high-gain 65-dB 2-stage amplifier with 61-degree phase margin and 38MHz unity-gain frequency is used. In order to mitigate voltage supply noise, an extra transistor is added (see top-right NFET in the Figure).
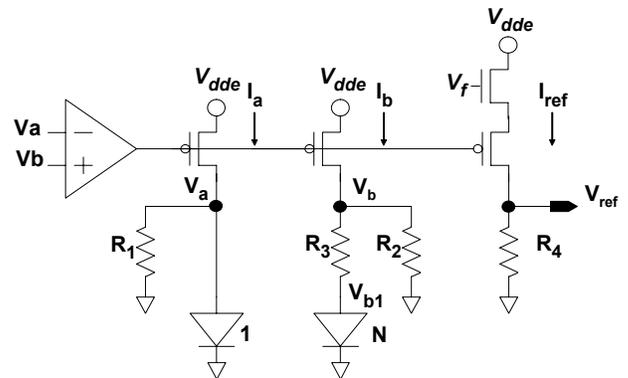
Figure 8   Sub-1V reference generation.

To separate flexible level generation and temperature behavior, I_{R3} and I_{R2} have opposing temperature coefficients, while R_4 is used to determine the reference level. The generated reference level can be expressed as

$$V_{ref} = \frac{R_4}{R_2}[V_a + \frac{R_2}{R_3}V_t \ln(\frac{N \cdot R_2}{R_1})] \quad (3)$$

Figure 9 shows the temperature dependence curve of the reference generator. The temperature coefficient is below 50ppm.
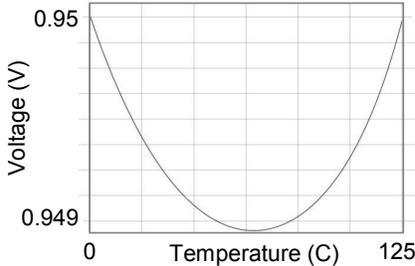
Figure 9   **Temperature behavior of reference.**

The current consumption is less than 20μA. Since the regulator itself consumes less than 100μA, the overall overhead is small. The area is 172x215 μm². At 0.95V, the overall transient response of the regulator at over 1 GHz switching remains under 20 mV.

6. METHODOLOGY

As shown in Figure 10, a concurrent mixed-signal design methodology is employed, where selective custom design, analog design, and digital logic design are performed in parallel. A competitive design schedule is achieved by attempting to mimic conventional synthesis-based ASIC methods whenever possible:

- Custom digital cells are created such that they are fully-compatible with the standard ASIC design methodology.

- Analog blocks are packaged as standard cores and thus can be treated as conventional blocks during link core integration.

- Core-level timing and functional verification are performed frequently to avoid re-design late in the design process.
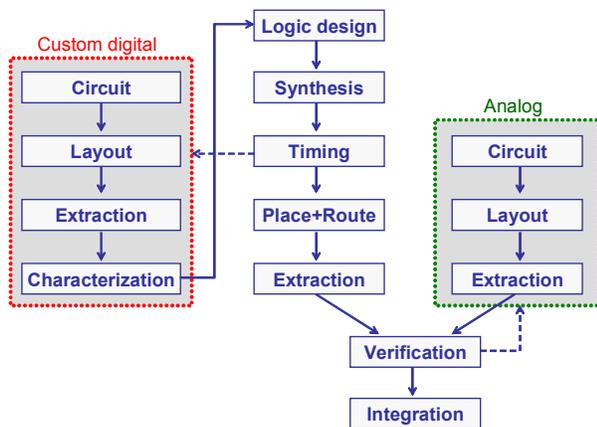
Figure 10   **Semi-custom design methodology.**

7. FULL MULTI-PROTOCOL RECEIVER CHIP

To verify the premises behind our approach, the ability to run large product-class multi-protocol multi-Gbps link blocks using under-1V on-chip regulation has been evaluated. To this end, a full 3.2-Gbps multi-protocol receiver logic chip has been developed, manufactured, and tested. A CMOS IBM 0.13μm technology with nominal 1.2V supply was used to manufacture the chip.

The chip, which is shown in Figure 11, benchmarks the presented approach versus a standard-voltage approach. The link receiver is fully self-testable by on-chip built-in self-test (BIST) logic to accomplish at-speed link performance evaluation. An on-chip PLL generates the necessary 1.6 GHz clocks.
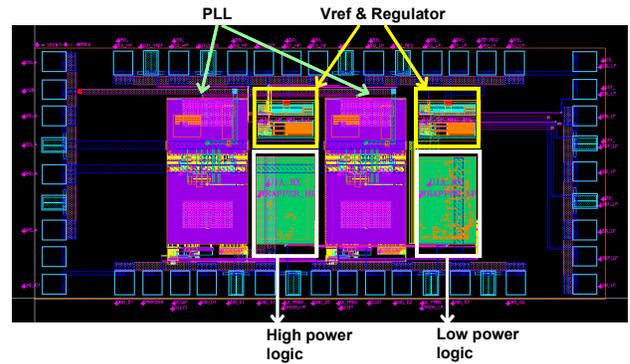
Figure 11   **Layout for manufactured chip.**

The right side of the chip includes regulator-powered semi-custom logic, where the regulator can be bypassed for stand-alone logic testing purposes. The left side features the conventional approach, together with a stand-alone regulator that can be tested independently.

Implementing the presented approach has an impact on design resources. To estimate such resource overhead, quantitative data was gathered during the design of this chip. Based on the resource effort and the characteristics of the latest design methodology, it is estimated that the incremental labor needed for this approach on a production-level link is only 0.8 person-years (PY). The major part of this incremental effort consisted of custom cell design, regulator optimization, and extra chip integration effort.

8. TESTING RESULTS

Theory and simulation predict an approximately linear behavior for power consumption versus regulator-supplied voltage. Figure 12 shows laboratory-measured results (room temperature) for the new regulator-powered receiver that indeed suggest a near-linear relationship.
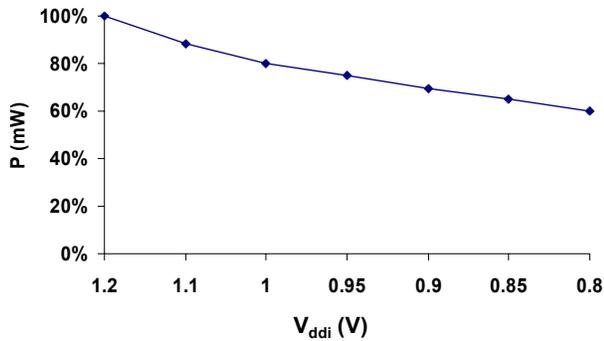
Figure 12   Normalized receiver logic power.

While functionality is correct at lower supplies than expected (near 0.8V) for the manufactured wafer, the power savings are also slightly lower than expected at each supply level. Adjusting for process effects and sub-optimal physical design, we estimate that realistic power savings for the presented approach amount to 25%-30% for this application.

9.   CONCLUSIONS

The work presented indicates that combining low supply voltages, an integrated voltage-island environment, and selective custom microprocessor design techniques result in an effective approach to reducing power while maintaining high performance, unmodified core interfaces, and a standard ASIC methodology.

The approach has been illustrated in detail on complex CMOS serial links that must meet multi-protocol specifications and present a simple, standard core interface. Power savings of over 25% is demonstrated for a 3000-gate multi-protocol receiver, while 3.2-Gbps performance is retained. The results suggest that the approach is effective for realistically complex links.

10.   REFERENCES

[1]   H. Banba, H. Shiga, A. Umezawa, T. Miyaba, T. Tanzawa, S. Atsumi, and K. Sakui, "A CMOS bandgap reference circuit with sub-1-V operation", IEEE Journal of Solid State Circuits, volume 34, issue 5, May 1999. Pages: 670 -674.

[2]   W.-H. Chen, G.-K. Dehang, J.-W. Chen, and S.-I. Liu, "A CMOS 400-Mb/s serial link for AS-memory systems using a PWM scheme", IEEE Journal of Solid State Circuits, volume 36 issue 10, Oct 2001. Pages: 1498 -1505.

[3]   T. Endoh, K. Sunaga, H. Sakuraba, and F. Masuoka, "An on-chip 96.5% current efficiency CMOS linear regulator using a flexible control technique of output current", IEEE Journal of Solid State Circuits, volume 36, issue 1, Jan 2001. Pages: 34 -39.

[4]   R. Farjad-Rad, C.-K.K. Yang, and M.A. Horowitz, "A 0.3-µm CMOS 8-Gigabit/s 4-PAM serial link transceiver", IEEE Journal of Solid State Circuits, volume 35, issue 5, May 2000. Pages: 757 -764.

[5]   C.-K. Ken Yang; R. Farjad-Rad, and M.A. Horowitz, "A 0.5µm CMOS 4Gb/s serial link transceiver with data recovery using oversampling", IEEE Journal of Solid State Circuits, 33-5, May 1998. Pages: 713 -722.

[6]   J.M. Khoury and K.R. Lakshmikumar, "High speed serial transceivers for data communication systems", IEEE Communications Magazine, volume 39, issue 7, July 2001. Pages: 160 -165.

[7]   M.-J.E. Lee, W.J. Dally, J.W. Poulton, P. Chiang, and S.E. Greenwood, "An 84-mW 4-Gb/s clock and data recovery circuit for serial link applications", Symposium on VLSI Circuits, 2001. Digest of Technical Papers, 2001. Pages: 149 -152.