High-Speed CMOS Dual-Modulus Prescalers for Frequency Synthesis

by

Ranganathan Desikachari

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented October 1, 2003
Commencement June 2004

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

HIGH-SPEED CMOS DUAL-MODULUS PRESCALERS FOR FREQUENCY SYNTHESIS

CHAPTER 1. INTRODUCTION

Frequency synthesizers are critical components for frequency translation and channel selection in wireless transceivers. Synthesizer design is a challenging task due to the stringent requirements imposed by RF systems.

This chapter provides an overview of the concepts of PLL-based frequency synthesis and the significance of variable-moduli prescalers.

1.1. Motivation

Oscillators and frequency synthesizers are key elements in a radio system that provide a controlled frequency source for receive signal down conversion and transmit signal up conversion. A stand-alone oscillator or a voltage controlled oscillator (VCO) does not have the required frequency stability to satisfy the low phase noise requirement. Therefore frequency synthesis is necessary to obtain accurate high frequencies from a precise low frequency crystal oscillator. Phase locked loops (PLLs) are often used to provide negative feedback in frequency synthesizers to suppress the phase noise due to oscillators.

Figure 1.1 shows a generic PLL-based synthesizer. The reference divider $\div$R can be used to scale highly accurate crystal based input frequencies down to desired levels for the PLL module. The PLL consists of a phase-frequency detector (PFD) and a loop filter (LF) apart from the VCO. The operation of the PLL and the programmable counter in the feedback path allow generation of accurate high frequencies from a pure

low frequency signal. The programmable P-counter is usually preceded by a prescaler ($\div$N) that scales down the high output frequencies to a range at which standard CMOS dividers can be implemented.

A dual-modulus division gives the flexibility to select channels on the basis of the number of times each of the moduli is selected. The dual-modulus prescaler could also be used in the feedback to obtain fractional output frequencies. Such synthesizer architectures, called fractional-N frequency synthesizers, allow the reference frequencies to be higher. PLL stability considerations require loop bandwidths to be a fraction of the reference frequency. So fractional-N synthesizers allow for higher loop bandwidths, resulting in faster output settling and lower oscillator noise characteristics.
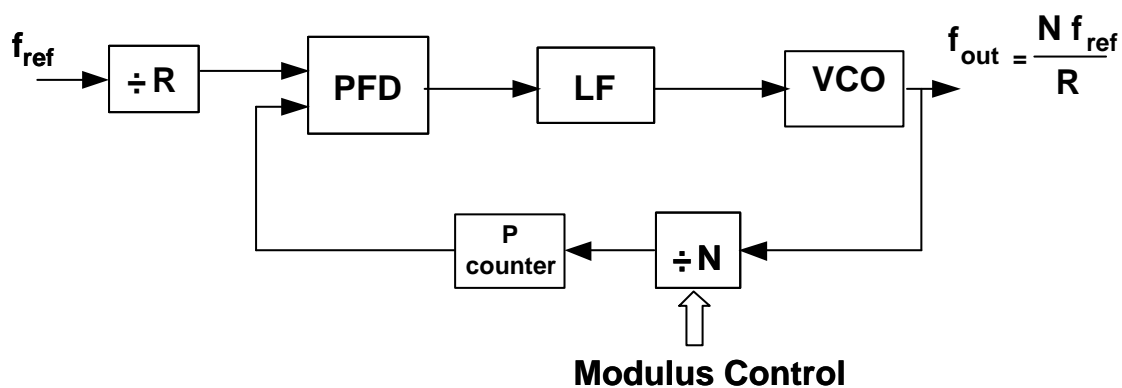
FIGURE 1.1: A typical integer-N PLL frequency synthesizer.

In several RF systems, the synthesizer is commonly partitioned into three separate chips: the VCO; the dual-modulus prescaler; and the channel selection logic(program counters and swallow counters) along with the phase/frequency detectors. As the VCO and the prescaler operate at a maximum frequency, they are usually fabricated in silicon bipolar or GaAs technologies and the rest in CMOS technology. Although these technologies support faster transistors, they are not very cost-effective. With the growing push towards integration of the entire synthesizer, research progress is being gradually made towards obtaining comparable performance with CMOS processes.

In order to tackle the multiple challenges presented by the requirements of wireless systems, it is necessary to understand the PLL frequency synthesizer at the system-level before analyzing and designing dual-modulus prescalers. These system level issues and the transistor-level implementation of dual-modulus prescalers have been addressed in this thesis.

1.2.   Thesis Organization

To provide a system level understanding of PLL-based frequency synthesizers, it is essential to discuss the features of commonly used structures. An overview of frequency synthesizer architectures and their characteristics, as well as a brief review of phase-locked loops are discussed in the Chapter 2. Chapter 3 deals with the details of the pulse-swallow topology for dual-modulus division. In Chapter 4 the design considerations and the actual implementation of the dual-modulus prescaler is discussed. The design challenges presented by the input/output buffers that interface this circuit to the outside world are also presented. Chapter 5 concludes the thesis with a compilation of the obtained simulation and measurement results.

CHAPTER 2. PLL-BASED FREQUENCY SYNTHESIZERS

With the boom of the wireless communications market, wireless transceivers have become ubiquitous. At the heart of every transceiver is a frequency synthesizer, required for accurate channel-selection. This chapter is an introduction to the concept of PLL-based frequency synthesis.

2.1. Introduction to Frequency Synthesizers

A transceiver (transmitter-receiver) is the building block that interfaces between the end user and the transmission medium. The transceiver consists of three blocks [1] - *the user-end* interfaces between the user information and the digital data representation; *the back-end* modulates and demodulates the digital data to and from the analog baseband signal that is suited for the transmission technique used (QPSK, GMSK etc.); and *the front-end* block that does the transmission, reception and frequency conversions.

Modern communication protocols always allocate closely located channels at very high frequencies. For example [2], Bluetooth, a short range wireless protocol, allocates 79 channels from 2.402 GHz to 2.480 GHz resulting in a 1MHz channel spacing. The phase noise from the LO needs to be low enough not to interfere with frequencies in the adjacent channels. However, a stand-alone oscillator with sufficiently high-Q, will not be tunable over a 79 MHz band. Additionally, even crystals do not have resonance frequencies as high as 2.4 GHz.

The above discussion illustrates the need for a block that synthesizes many discrete frequencies from one or more fixed reference frequencies. Figure 2.1 illustrates this in the context of a generic transceiver.
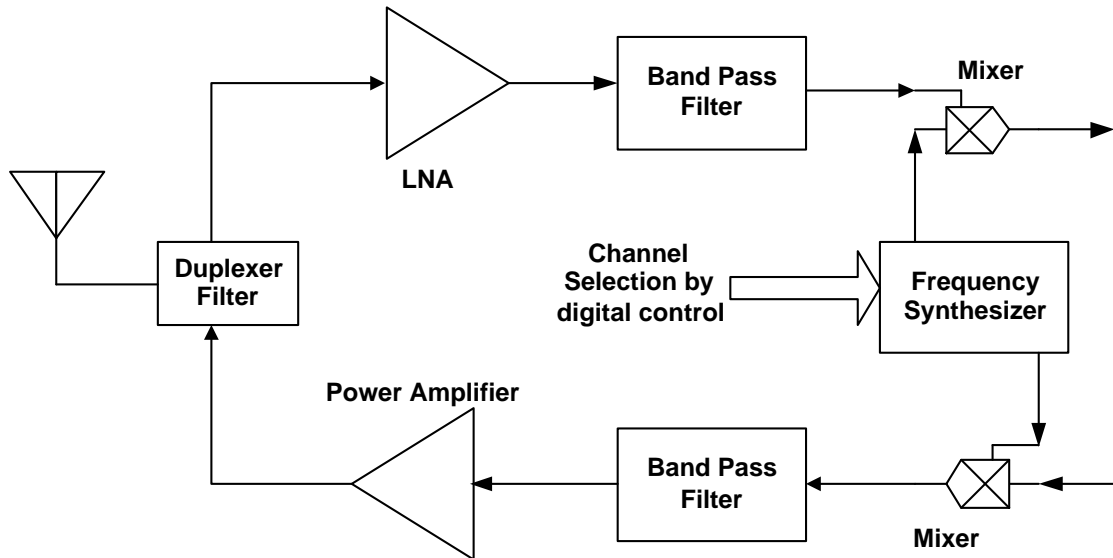
FIGURE 2.1: Frequency synthesizer operation in a generic transceiver.

As many discrete frequencies need to be generated it is impractical to have a reference frequency for each. Ideally, the reference frequency is a single spectrally pure frequency, typically generated from a piezo-electric crystal. This leads to the idea of a control input to generate required frequencies from a single reference.

Phase Locked Loops (PLLs) are negative feedback systems whose output frequency can be digitally controlled with the help of a precise clock at its input as reference. The reference phase noise is suppressed within the loop bandwidth by the negative feedback. High output frequencies can be obtained from accurate references by frequency division in the feedback path. Therefore they are ideal devices for frequency synthesis. Figure 2.2 shows a typical PLL-based implementation of frequency synthesizer. This is also referred to as an integer-N synthesizer, discussed in detail in later sections.

$$f_{out} = \frac{N f_{ref}}{R}$$
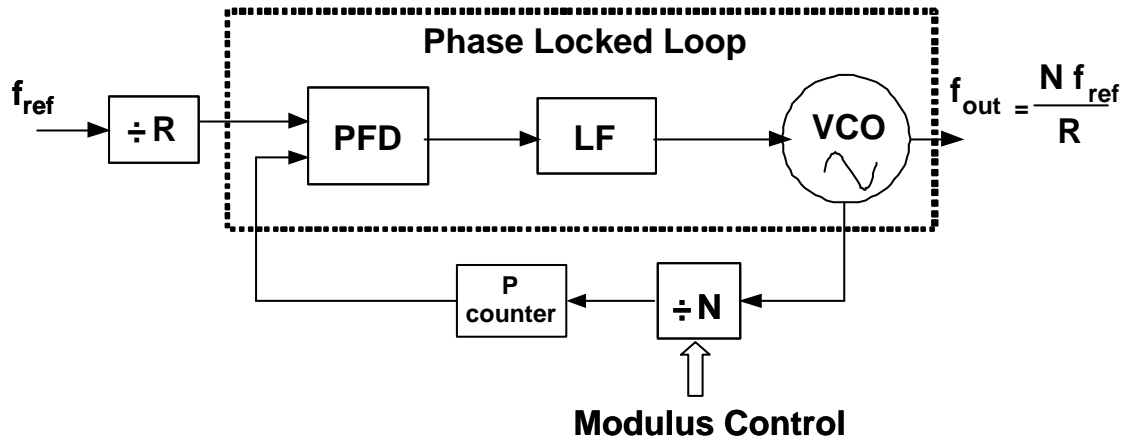
FIGURE 2.2: PLL-based frequency synthesizer.

2.2. Characterization of Frequency Synthesizers

Modern wireless standards impose several stringent requirements which challenge the design of frequency synthesizers. This is exemplified by the Bluetooth protocol mentioned in the previous section. Some of the important parameters that characterize the performance of synthesizers are highlighted in this section.

2.2.1. Frequency Range

The range of frequencies generated by the synthesizer is defined by the wireless standard (900 MHz, 1.9 GHz, 2.4 GHz etc.). In most cases, the frequency must be varied in small increments determined by the channel spacing. This frequency resolution could be as low as 30 kHz. A specification on the output frequency accuracy, as well as on the channel's upper and lower edges requires the error of the synthesizer to be less than a few parts per million. The generated output frequencies could experience short term (drift) or long term variations (aging) due to the environment. So a frequency stability specification is usually defined with respect to time, temperature, power supply etc.

2.2.2. Spectral Purity

The spectral purity of the synthesized output can deviate due to the nonidealities in the components of the PLL. The ideal output spectrum of a frequency synthesizer should be a single tone at the desired frequency in order to provide the reference frequency for frequency translation. A single tone in the frequency domain is equivalent to a pure sinusoidal waveform in the time domain. The random and systematic amplitude and phase deviations from the desired values produce energy in frequencies other than the desired frequency. When this energy is mixed with the received RF signal or modulated baseband signal, undesired sidebands are created. Phase noise and spurious tones are the two key parameters to measure the quality of a frequency synthesizer. In this section, the effects of phase noise and spurious tones on a transceiver have been investigated.

Phase noise is the phenomenon of phase disturbance of oscillators and has been modeled and described extensively in literature[3, 4, 5, 6]. The ideal synthesizer has a pure sinusoidal waveform as given by Eq. (2.1).

$$v(t) = V_0 Cos(2\pi f_0 t) \tag{2.1}$$

When amplitude and phase fluctuations are included, the waveform becomes

$$v(t) = [V_0 + \epsilon(t)]Cos[2\pi f_0 t + \phi(t)] \tag{2.2}$$

where $\epsilon(t)$ represents an amplitude fluctuation and $\phi(t)$ represents a phase fluctuation. Because amplitude fluctuations can be removed or greatly reduced by a limiter, phase modulation is a bigger concern in frequency synthesizer design. The phase fluctuations could arise in three different ways: *systematic variations*, due to aging of the resonator material; *deterministic periodic variations* due to unwanted phase or frequency modulations in the PLL, and *random variations* due to noise sources such as thermal, shot or flicker noise in the devices.

The phase noise of oscillators or synthesizers is measured as the ratio of the noise

power in a 1Hz bandwidth at a certain offset frequency from the carrier to the noise power of the carrier.

$$\phi(f) = 10 \ log\frac{P_{noise}}{P_{carrier}}(dBc/Hz) \tag{2.3}$$

Although the loop components suffer from the above mentioned noise sources, the two important contributors of phase noise in a PLL are the input reference noise and the VCO noise. The PLL loop bandwidth is a design parameter that is determined on the basis of the dominant noise source. A linear model of PLLs has been discussed in a later section.

As shown in Figure 2.3, the sidebands caused by the phase modulation appears as a *phase noise skirt.* Sometimes energy is concentrated at frequencies other than the desired frequency, resulting in **spurious tones** appearing as spikes above the phase noise skirt. These tones are usually artifacts of reference frequency feed-through due to charge-pump nonidealities or due to any periodicity introduced by the modulus-selection operation. Some standard strategies used to alleviate the problem of spurious tones are to use larger loop filter capacitors, notch filters to suppress reference feed-through and the use of delta-sigma modulators to shape the noise due to modulus-selection out of the band of interest [7, 8].

2.2.3. Transient Response Requirements

As shown in the phase-locked architecture of Figure 2.2, the modulus variation by control signal would result in a loop transient. Every time a different channel is selected, the PLL needs to lock to the new frequency. The lock time of synthesizers is especially a critical parameter in fast frequency-hopped spread-spectrum systems.

An interesting analysis to model the loop settling behavior of PLL architectures has been carried out in [1]. The modulus change in the system can be modeled with
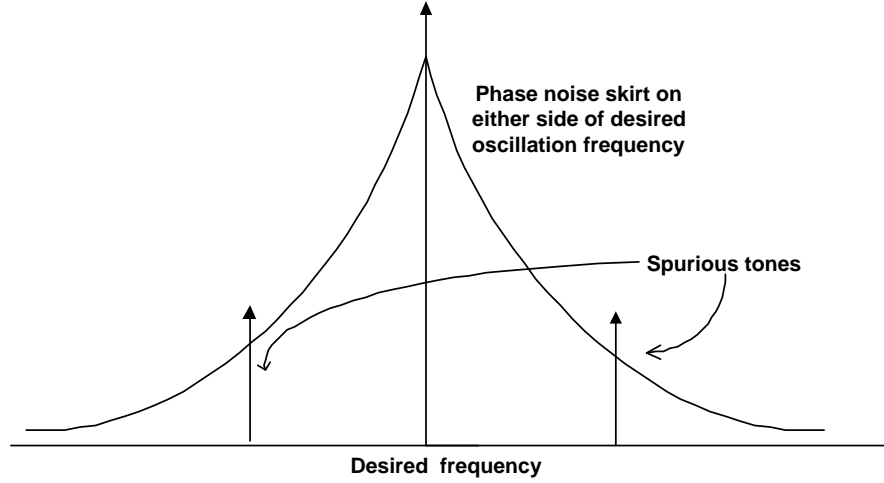
FIGURE 2.3: Phase noise and spurious tones.

a simple feedback topology as illustrated in Figure 2.4. Suppose the divider modulus (N) variation corresponds to a small step ϵ in the feedback factor, the closed-loop jitter transfer function of this system is equal to

$$
\begin{aligned}
Y(s) &= \frac{H(s)}{1 + (N + \epsilon)H(s)} X(s) \\
&\approx \frac{H(s)}{1 + NH(s)} \cdot \frac{1}{1 + \epsilon/N} X(s) \\
&\approx \frac{H(s)}{1 + NH(s)} (1 - \frac{\epsilon}{N}) X(s)
\end{aligned}
\tag{2.4}
$$

where, H(s) is the phase-domain transfer function of the PLL. The above relation implies that the modulus change is equivalent to multiplying the input by (1- ϵ/N). The output frequency's response can thus be estimated with conventional second-order transient equations to a step input.

2.3. PLL System Description

Since the frequency synthesizers are based on the phase-locking principle, a typical PLL system has been described in this section. PLLs are negative feedback systems that
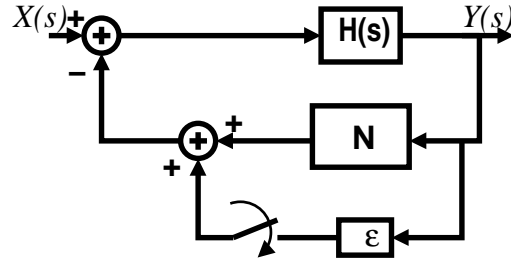
FIGURE 2.4: Modulus switching transient model

operate on excess phase of nominally periodic signals. Their function is to lock the frequencies and phase of its two input signals with as small an error as possible.

2.3.1. Basic Operation of PLL

The simplest PLL, shown in Figure 2.5(a) consists of a phase detector (PD), low-pass filter (LPF) and a voltage-controlled oscillator (VCO). The PD serves as an error amplifier in the feedback loop, and tries to minimize the phase difference $\Delta\Phi$ between X(t) and Y(t). The loop is considered "locked" if $\Delta\Phi$ is constant with time.
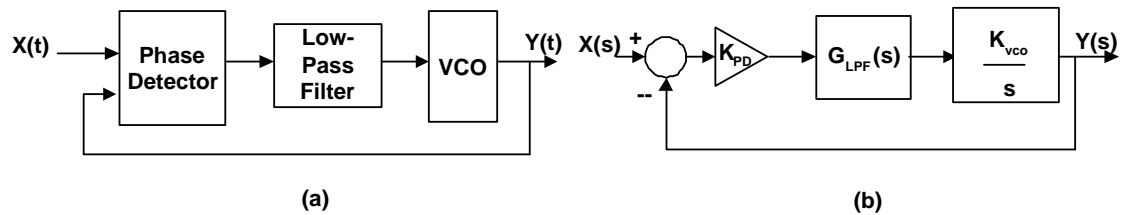
FIGURE 2.5: Phase-locked loop (a) block diagram,(b) linear model.

In the locked condition, the PD produces an output proportional to the phase error between the input and output signals. The LPF averages out this error so that a voltage error ΔV corresponding to the phase error is built up at the input of the VCO. The output frequency is modulated to minimize the phase and frequency error.

2.3.2. PLL Loop Dynamics

Although the PLL transient is a nonlinear process, a linear approximation can be used to arrive at a phase-domain PLL model that gives intuitive insight into the tradeoffs involved in the design. From the linear PLL model of Figure 2.5(b), the closed-loop transfer function, also known as the *jitter transfer function*, is given by,

$$
\begin{aligned}
H(s) &= \frac{\Phi_{out}(s)}{\Phi_{in}(s)} \\
&= \frac{K_{PD}K_{VCO}G_{LPF}(s)}{s + K_{PD}K_{VCO}G_{LPF}(s)}
\end{aligned}
\tag{2.5}
$$

If $G_{LPF}(\mathrm{s}) = \frac{1}{1+\frac{s}{\omega_{LPF}}}$,

$$
H(s) = \frac{K_{PD}K_{VCO}}{\frac{s^2}{\omega_{LPF}} + s + K_{PD}K_{VCO}}
\tag{2.6}
$$

From the above equations, it can be shown[1, 3] that the finite phase error is inversely proportional to the loop gain $K_{PD}K_{VCO}$. Further, the above simple phase detector and low pass-filter combination does not give independent control of the loop-bandwidth and damping factor. These limitations necessitate the use of *charge-pump PLLs (CPLLs)*. A typical CPLL schematic is shown in Figure 2.6.
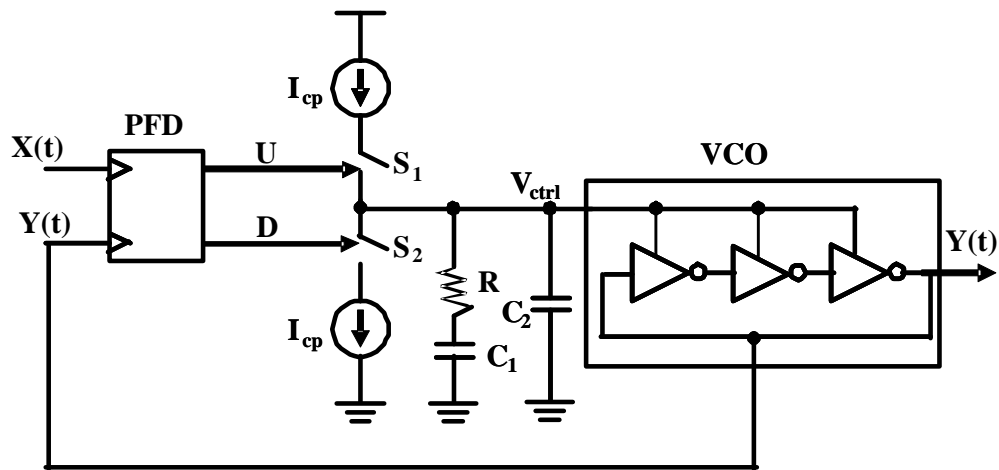
FIGURE 2.6: Charge-Pump Phase-Locked Loop

This CPLL structure has been analyzed extensively in the literature [9, 10]. The salient features of this PLL structure have been discussed below without digressing into the mathematical details of the system.

The simple phase detector has been replaced by a phase-frequency detector (PFD). The PFD is a circuit that can detect both phase and frequency difference between the input and output signals. This ability is a consequence of realizing the PFD using digital sequential logic that compares clock edges instead of continuous input-output phase comparisons. In order to convert the phase error pulses into a control voltage, charge-pumps are used to charge up/down the loop filter depending on the state of the PFD control signals. By virtue of incorporating a charge-pump and loop filter to average the phase error pulses, the CPLL has ideally infinite DC gain, making the phase error negligibly small. The loop filter capacitance C_1 and the VCO, each contribute one pole at the origin in the jitter transfer function. The resistance is, therefore, introduced to compensate for the net phase shift around the loop. This implementation of the filter gives the additional advantage of independent control of the loop-bandwidth and damping factor [1].

Most PLL-based synthesizers use this basic topology with several variations in the design of the individual building blocks. The switching introduced by the digital blocks and the charge pumps can affect the purity of frequency synthesizer outputs. This has been discussed in later sections.

2.4. Frequency Synthesizer Architectures

Output frequencies of frequency synthesizers vary in discrete steps corresponding to the channel spacing i.e. $f_{out} = f_o + \text{k} \cdot f_{ch}$ where f_o is the lower limit of frequency. It was shown in Figure 2.1 that the 'k' is selected by digital control. The need for output frequencies to be accurately locked to a particular channel mandates the use of a PLL. Some popular architectures for frequency synthesis and their features are discussed here.

2.4.1. Static-Moduli Frequency Synthesizers

The classic PLL frequency synthesizer (Figure 2.7(a)) comprises a reference oscillator and two static modulus dividers so that,

$$\frac{f_{out}}{N} = \frac{f_{ref}}{R} \Rightarrow f_{out} = \frac{N}{R} \cdot f_{ref} \tag{2.7}$$

Thus, by varying the moduli, different output frequencies can be synthesized. Output frequencies can be incremented in steps of f_{ref}/R , and this is the frequency at which the PFD is updated. As mentioned earlier, this requires the PLL loop bandwidth to be a fraction of this update rate. The lower bandwidth is a big disadvantage in terms of the reference phase noise suppression and slower settling times.
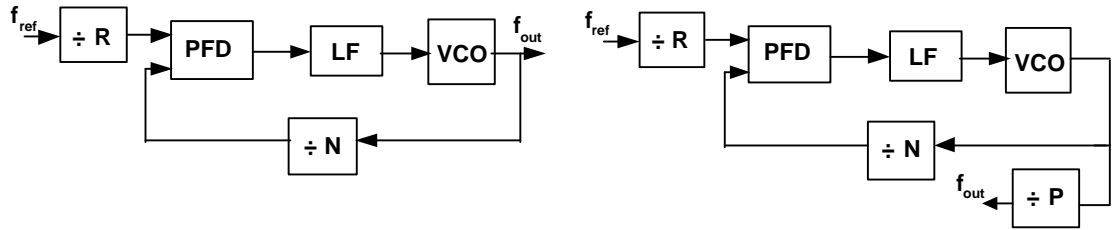
FIGURE 2.7: (a) Classical static modulus divider; (b) Modified static modulus divider

A modification made to the above architecture is to include a static divider following the PLL as shown in Figure 2.7(b), so that the output frequencies are given by $f_{out} = \frac{N}{R \cdot P} \cdot f_{ref}$ but the PLL update rate is f_{ref}/N , a factor of P improvement.

However, although this modification improves the PLL loop bandwidth constraint, the operating frequency of the PLL and the $\div N$ need to be higher.

2.4.2.   Integer-N Frequency Synthesizers

The integer-N frequency synthesizer is probably the most widely used architecture. The feedback divider now consists of a static divider as well as a dual-modulus prescaler coupled with two counters. The two counters are referred to as the **swallow-counter** and **program counter**. The swallow counter, or the channel-spacing counter, can be programmed to enable channel selection. The prescaler divides by N+1 until the swallow counter overflows after which the overflow bit will set the prescaler in divide-by-N mode until the program counter overflows. Stated more explicitly, the program counter (also known as the frame counter) determines the total number of VCO cycles required for the above operation. The detailed operation and the math involved in an integer-N system to realize channel selection on the basis of programmed counters has been discussed in later chapters.

2.4.3.   Fractional-N Synthesizers

As mentioned above, integer-N PLLs are restricted in reference frequencies by the channel spacing. The principle of *pulse-swallow* or pulse-removal can be used to implement fractional division ratios. Depicted in Figure 2.8, are two implementation strategies for fractional-N synthesis. Figure 2.8(a) incorporates a pulse-remover which blocks one VCO pulse upon assertion of the "remove" control. The average locked frequency is hence,

$$f_{out} = f_{ref} + 1/T_p \tag{2.8}$$

where, T_p is the period with which the pulse-remove command is applied. Modern implementations however are based on dithering between two moduli. If the prescaler divides by N for A VCO pulses and by N+1 for B pulses, the average divide ratio would be equal to
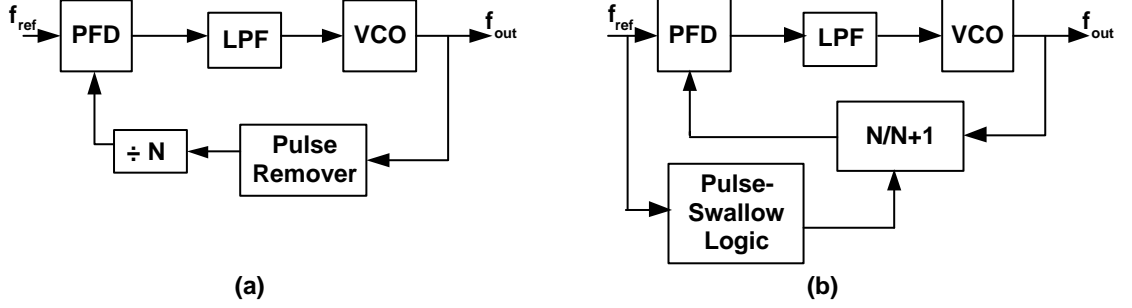
FIGURE 2.8: Fractional-N synthesis based on (a) pulse removal, (b) dual-modulus prescaler

$$N_{avg} = \frac{NA + (N+1)B}{A+B} = N + \frac{B}{A+B} \tag{2.9}$$

$$f_{average} = \frac{A+B}{\frac{A}{N} + \frac{B}{N+1}} \tag{2.10}$$

$$\Rightarrow f_{average} = N \cdot f \tag{2.11}$$

where, N represents the integer potion of the division modulus and f represents the fractional portion.

The obvious advantage of the above scheme is that the channel spacing need only be a fraction of the reference frequency, allowing for higher PLL loop bandwidths. the critical drawback of fractional-N synthesis is that the periodicity introduced by the modulus-select operation usually appears as sidebands or spurs. The conventional technique to eliminate the spurs is by dithering the modulus. To avoid an increase of the noise floor, delta-sigma techniques can be used to shape the noise out of band which can then be filtered out. This randomization of the modulus control to suppress spurs has been discussed in [1].

CHAPTER 3. DUAL-MODULUS PRESCALERS

The applications of dual-mod prescalers has been highlighted earlier in the context of integer-N and fractional-N synthesizers. An analysis and a discussion of their operation are carried out in this chapter.

3.1. Dual-Modulus Operation

The complexity of the N counter in PLL frequency synthesizers has grown over the years. In addition to a straightforward N counter, it has evolved to include a prescaler. This structure, illustrated earlier in Figure 2.2, has developed as a solution to the problems inherent in using the basic divide-by-N structure to feed back to the phase detector when very high-frequency outputs are required. For example, suppose a 900 MHz output is required with 10 kHz channel spacing. A 10 MHz reference frequency might be used, with the R-divider set at 1000. Then, the N-value in the feedback would need to be of the order of 90000. This would need a 17-bit counter capable of dealing with an input frequency of 900 MHz.

To handle this range, it makes sense to precede the programmable counter with a fixed counter element (the prescaler), to bring the very high input frequency down to a range at which standard CMOS will operate. However, using a standard prescaler introduces other complications. The system resolution, or the effective channel spacing, is now degraded by P, the modulus of the prescaler. This issue can be addressed by using a dual-modulus prescaler (Figure 3.1). It has the advantages of the standard prescaler but without any loss in system resolution. A dual-modulus prescaler is a counter whose division ratio can be switched from one value to another by an external control signal. By using the dual-modulus prescaler with an 'S' and 'P' counter one can still maintain

an output resolution specified by the input to the PLL (F_1).
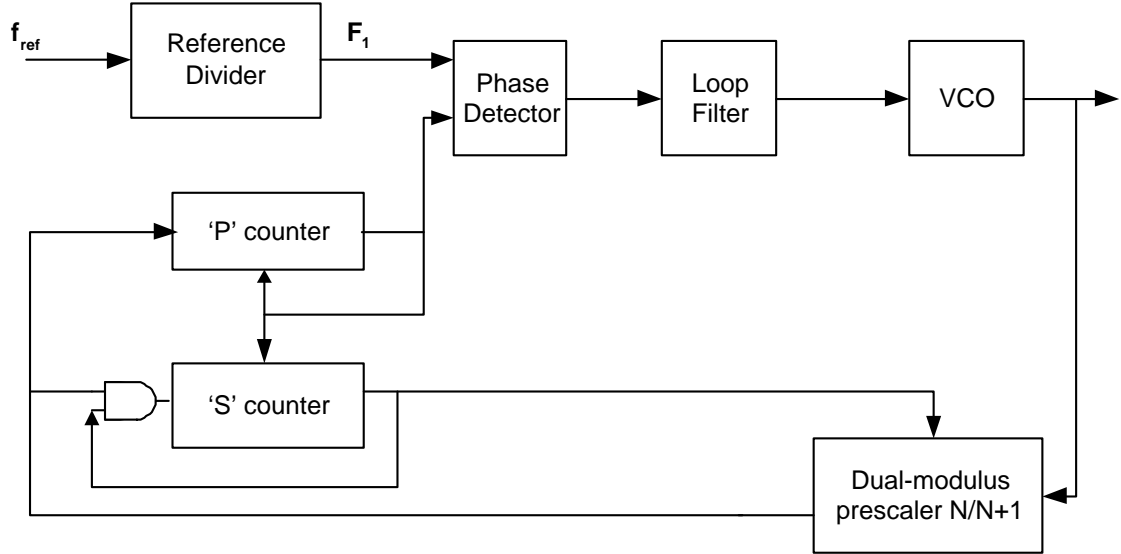
FIGURE 3.1: Pulse-swallow integer-N frequency synthesizer.

As long as the S counter has not timed out, the prescaler divides down by N + 1. So, both the S and P counters will count down by 1 every time the prescaler counts (N + 1) VCO cycles. This means the S counter will time out after ((N + 1) S) VCO cycles. At this point the prescaler is switched to divide-by-N mode. The P counter still has (P - S) cycles to go before it times out. So after ((P - S) N) more cycles, the system is now reset to the initial condition. Expressing the above discussion mathematically, total number of VCO cycles for one dual-modulus division, is

$$
\begin{aligned}
M &= (S \times (N + 1) + (P - S) \times N) \\
&= (SN + S + PN - SN) \\
\implies M &= (S + PN)
\end{aligned}
\tag{3.1}
$$

Typically the S counter is called the *swallow counter*, and the P counter is the *program counter*.

Consider the expression M = S + PN. To ensure a continuous integer spacing for M, S must be in the range 0 to (N - 1). Then, every time P is incremented there is

enough resolution to fill in all the integer values between PN and (P + 1)N. As was already noted for the dual-modulus prescaler, P must be greater than or equal to S for the dual modulus prescaler to work. From these we can say that the smallest division ratio possible while being able to increment in discrete integer steps is:

$$
\begin{aligned}
M_{MIN} &= (P_{min} \cdot N) + S_{min} \\
&= ((N - 1) \cdot N) + 0 \\
&= (N^2 - N)
\end{aligned}
\tag{3.2}
$$

The highest value of M is given by

$$
M_{MAX} = (P_{max} \cdot N) + S_{max}
\tag{3.3}
$$

In this case S_{max} and P_{max} are simply determined by the size of the S and P counters. The range from M_{MIN} to M_{MAX} defines the multiple moduli of division.

3.2. Pulse-Swallow Architecture

One important factor that has not been addressed yet is how the dual-modulus operation can be implemented. The conventional implementation of any divider/prescaler is using digital counters. The division factor that can be easily realized using such logic is of the form 2^N, i.e. the pattern in which the counter counts repeats every 2^N cycles. To implement $2^N + 1$, therefore, one extra state of the system needs to be inserted over a single pulse duration in the repetitive pattern. This is referred to as a "pulse-swallow" operation.

The principle of operation of the pulse-swallow architecture can be explained by means of a simple divide-by-2/3 circuit (Figure 3.2). Figure 3.2(a) shows the simplest divider, a ÷2 implemented with a D flip-flop (DFF2). Now if another DFF (DFF1) and the combinational gate G are inserted in the feedback path of the divider(Figure 3.2(b)),
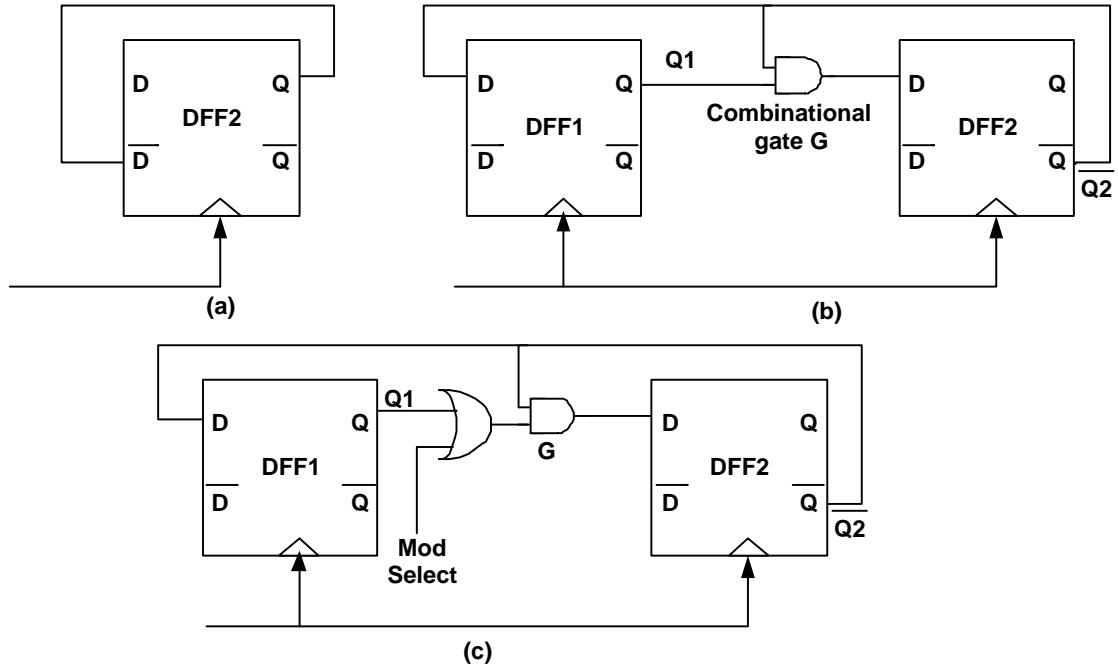
FIGURE 3.2: 2/3 dual-modulus divider(a) divide-by-2 ;(b) divide-by-3 circuit ;(c) 2/3 dual-modulus divider.

then the system can be in three states : $Q_1 \ \overline{Q_2} = 01$, 10, 11. The $Q_1 \ \overline{Q_2} = 00$ is obviously illegal as that implies the previous values of $\overline{Q_2}$ and G would have had to be in the impossible states of '0' and '1', respectively. In order to control the mod-select an extra gate is required, such as the OR gate in Figure 3.2(c). This simple 2/3 divider works in divide-by-2 mode when Mod Select is '1' and in divide-by-3 mode for Mod Select '0'. The above discussion can be extended to higher division moduli $2^N/2^{N+1}$ prescalers easily.

3.3. Technology Comparison - Bipolar Vs CMOS

The two most important performance parameters to be optimized in the design of prescalers are speed and power. The biggest limiting factor in this optimization is the technology. As mentioned earlier, the prescaler being one of the components working at full speed, is often implemented with bipolar or SiGe/GaAs technologies [11, 12, 13, 14].

One of the perennial questions that has been discussed often in several conference panel discussions and by RF engineers is the wisdom in pursuing RF-CMOS. This section compares bipolar and CMOS technology for RF applications, throwing light on some of their merits and demerits.

The key transistor figures of merit for RF and microwave applications are the unity-gain frequency f_T and the maximum power gain frequency f_{max}, 1/f noise corner frequency etc. A comparison of the figures of merits of several technologies suitable for wireless LAN applications has been tabulated in [15]. As silicon technologies are less expensive and more integrable, they would be the clear choice. However, it is not obvious as to which silicon technology ought to be used, bipolar/BiCMOS or CMOS. Some of the issues considered are listed below.

Device Performance Comparison: Bipolar transistors do have a lot of performance advantages over MOSFETs in RF and analog applications. Some of the important comparisons are:

- g_m/I ratio of bipolar transistors is always higher than MOSFETs [16]. The NPN transistor possess a higher inherent gain compared to NMOSFETs and hence, a higher drive capability.

- Bipolar transistors have lower 1/f noise than MOSFETs due to the absence of surface charge effects. This is a significant advantage for low-noise RF circuits.

- Bipolar transistors are sometimes considered to be modeled better than MOSFETs, especially in the deep-submicron processes currently used [17].

- Bipolar devices exhibit better device matching on the same die.

- Bipolar transistors, however, are more non-linear than MOS devices due to their exponential I-V characteristics. This is especially significant in the context of devices that are used as switches. However distortion introduced due to back-gate effects are absent in bipolar transistors.

- MOS devices have the advantage of the availability of complementary PMOS devices.

For RF circuits, bipolar devices do seem to possess more desirable features. Several commercial analog products, however, utilize the advantages of both the bipolar and MOS characteristics in BiCMOS processes.

Availability and Accessibility: It has been well established that CMOS is the most available and accessible process among all semiconductor technologies. Several foundries around the world offer a wide range of CMOS processes. This is the biggest advantage of RF CMOS over bipolar and BiCMOS processes.

Cost, Yield and Integration Levels: The main appeal of CMOS is the relatively low cost combined with high levels of integration. Shrinking device sizes is an attractive feature for digital CMOS circuits as it improves both speed and power dissipation [18]. This is an added motivation for CMOS integration of analog components. Bipolar transistors are also more prone to defect density as they are minority carrier devices. CMOS processes therefore have a higher yield. However, arguments presented in [17] point out that when costs associated with packaging and testing are included, the price tag on RF chips are not significantly different. Also, any RF or analog CMOS process requires more masks for good passive components, adding to fabrication costs. Combined with the performance advantages of bipolar devices, the cost factor advantage of CMOS can be challenged .

Based on above discussions it can be concluded that the choice of technology depends on the kind of application. If integration levels of wireless systems become sufficiently high for Radio on Chip (ROC) to be feasible, CMOS processes may reduce the entire chip-set into one big chip with small supporting chips. When the die cost of the chip is a significant portion of the overall system costs, CMOS could have significant edge. For several (low end) radio systems CMOS RF performance may be comparable to BiCMOS implementations and may be preferable. The attempt of this thesis is to realize high-speed, low-power dual-modulus prescalers in RF CMOS technology. The

implementation of the dual-modulus prescalers in Chapter 4 highlights the tradeoffs involved with the CMOS design.

3.4.   Current Mode Operation

CMOS static logic is widely used in mixed-signal integrated circuits because of its ease of design, high packing densities, wide noise margins, etc. The most significant feature is that the static power dissipation is nearly zero. However, its power dissipation at high frequencies due to the displacement current $C_{out}(dV_{out}/dt)$ accounts for dynamic power $P_{dynamic} \approx C_{out} (\Delta V_L)^2 f$. As illustrated in Figure 3.3, the current spikes during switching could flow through parasitic resistances and inductances associated with the V_{dd} and Gnd power supply grid networks, bond-pad, package parasitics, etc. and cause V_{dd} *bounce* or *Gnd bounce* by virtue of the $I \cdot R$ or $L \cdot \frac{dI}{dt}$ voltage drops. This kind of digital switching noise could show up as annoying glitches in the analog part of a mixed-signal chip. Although there has been much research progress in the modeling of this substrate coupling in mixed-signal ICs, the effects of the digital switching is difficult to predict, making it difficult to eliminate with conventional circuit and layout techniques [19].

CMOS static logic belongs to the category of voltage switching circuits in which V_{dd} or V_{ss} is switched to the output node. The fundamental reason for the the digital switching noise is that the power supply current is not held constant during output voltage transitions [20]. This observation has motivated the development of *source-coupled logic* circuits. Source-coupled logic circuits, also referred to as MOS current mode logic circuits (MCML) work on the principle of current steering controlled by an input to a differential pair. As shown in Figure 3.4, the tail current is steered on either side of the source-coupled pair and the output differential voltage determined by the tail current and load resistances. This kind of differential logic has several advantages over conventional CMOS as discussed in detail in this section.

(a)

(b)
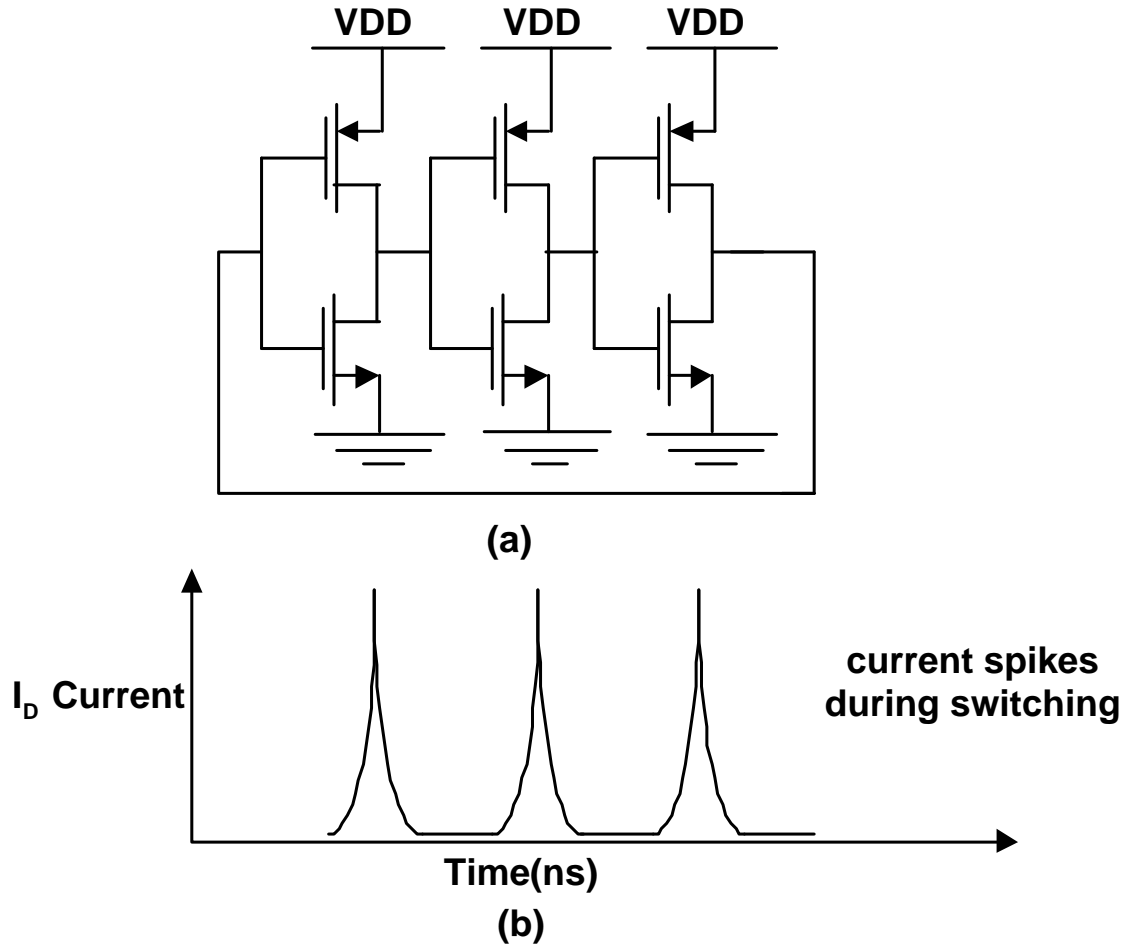
FIGURE 3.3: (a) CMOS ring oscillator;(b) Switching spikes in a CMOS inverter.

3.4.1. Speed-Power Advantage

Several comparisons of CMOS and MCML circuits have been carried out in literature [21, 22, 23]. Suppose a linear chain of N identical gates, all with an identical load capacitance C on each output node was compared and contrasted in the two different cases, the total propagation delay (D) of the chain of gates will be proportional to:

$$D_{CMOS} = \frac{N \times C \times V_{dd}}{0.5k \times (V_{dd} - V_t)^\alpha} \tag{3.4}$$

where, k and α are parameters depending on transistor dimension and process. Assuming the CMOS logic is clocked at a frequency equal to the inverse of the propagation delay,
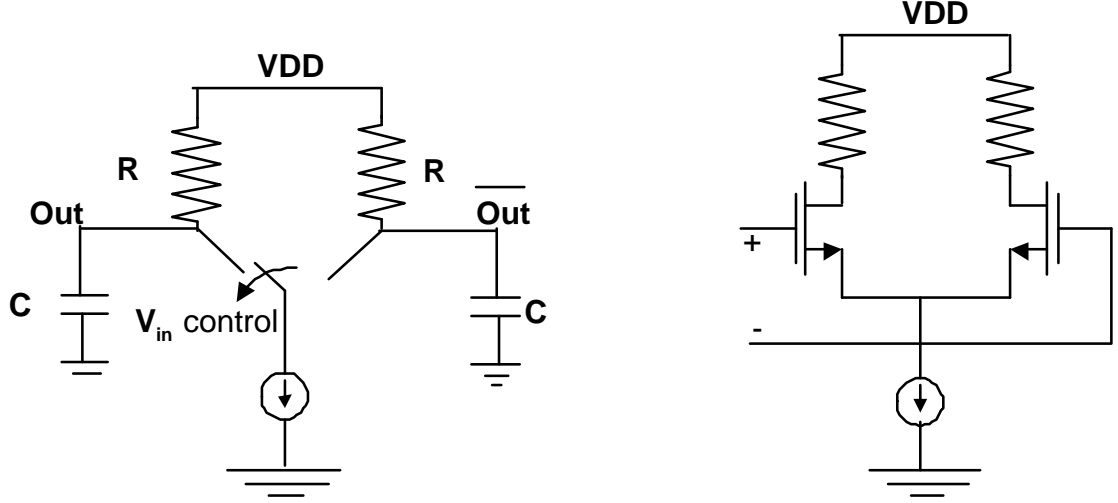
FIGURE 3.4: Principle of current-mode logic.

the dynamic power dissipation, power-delay and energy-delay products are given by :

$$P_{CMOS} \;=\; N \times C \times V_{dd}^2 \times \frac{1}{D_{CMOS}} \tag{3.5}$$

$$PD_{CMOS} \;=\; N \times C \times V_{dd}^2 \tag{3.6}$$

$$ED_{CMOS} \;=\; N^2 \cdot 2\frac{C^2}{k}\frac{V_{dd}^2}{(V_{dd}-V_t)^\alpha} \tag{3.7}$$

The objective of digital design is to optimize the energy-delay (ED) product. It can be derived that the optimized supply voltage for minimizing the ED product for CMOS is

$$V_{dd} = \frac{2V_t}{3-\alpha} \tag{3.8}$$

The power-delay equations for a CML inverter cascade are [22] :

$$D_{CML} \;=\; NRC = \frac{N \times C \times \Delta V}{I} \tag{3.9}$$

$$P_{CML} \;=\; N \times I \times V_{dd} \tag{3.10}$$

$$PD_{CML} \;=\; NIV_{dd} \times \frac{NC\Delta V}{I} = N^2 \times C \times \Delta V \times V_{dd} \tag{3.11}$$

$$ED_{CML} \;=\; N^2 C V_{dd}(\Delta V) \times \frac{NC\Delta V}{I} = \frac{N^3 C^2 V_{dd}\Delta V^2}{I} \tag{3.12}$$

$$\tag{3.13}$$

where, ΔV is the output voltage swing $= I \cdot R$

The above results indicate that CML circuits can be optimized by reducing the supply voltage, or the signal voltage swing, and by increasing the tail current.

Intuitively, the higher speeds of current-mode operation can be attributed to two main aspects - the transistors need not be completely turned on/off as in the case of CMOS, and the lower voltage swings can charge/discharge the output node capacitance much faster. The conventional power advantage of CMOS circuits does not hold at such high frequencies as their dynamic power dissipation is comparable to or even worse than the static power loss in CML circuits.

3.4.2. Common-Mode Noise Suppression

One of the most significant drawbacks of CMOS logic is the effect of the current spikes during switching. The large transient currents could lead to $L\frac{dI}{dt}$ voltage drops of the order of about 200 mV. Since many analog signals could be much smaller than this, such variations could be disastrous. The constant current drawn by source-coupled pairs reduces this noise coupling by a large extent.

3.4.3. Substrate Coupling

Another source of switching noise is the injection of currents into the substrate by charging/discharging of the drain-bulk capacitance (Figure 3.5(a)). In case of single-ended rail-to-rail CMOS logic, the voltage variation modulates the depletion widths causing a current

$$i_{sub} = C_{db}\frac{dv_{out}}{dt} \tag{3.14}$$

The use of differential logic in CML circuits cancels these substrate currents to a first order as illustrated in Figure 3.5(b). The total substrate current is now

$$i_{sub} = C_{db1}\frac{dv_{out}}{dt} + C_{db2}\frac{\overline{dv_{out}}}{dt} \tag{3.15}$$

The cancelation is not exact as C_{db} is non-linear and depends on the voltage across it.
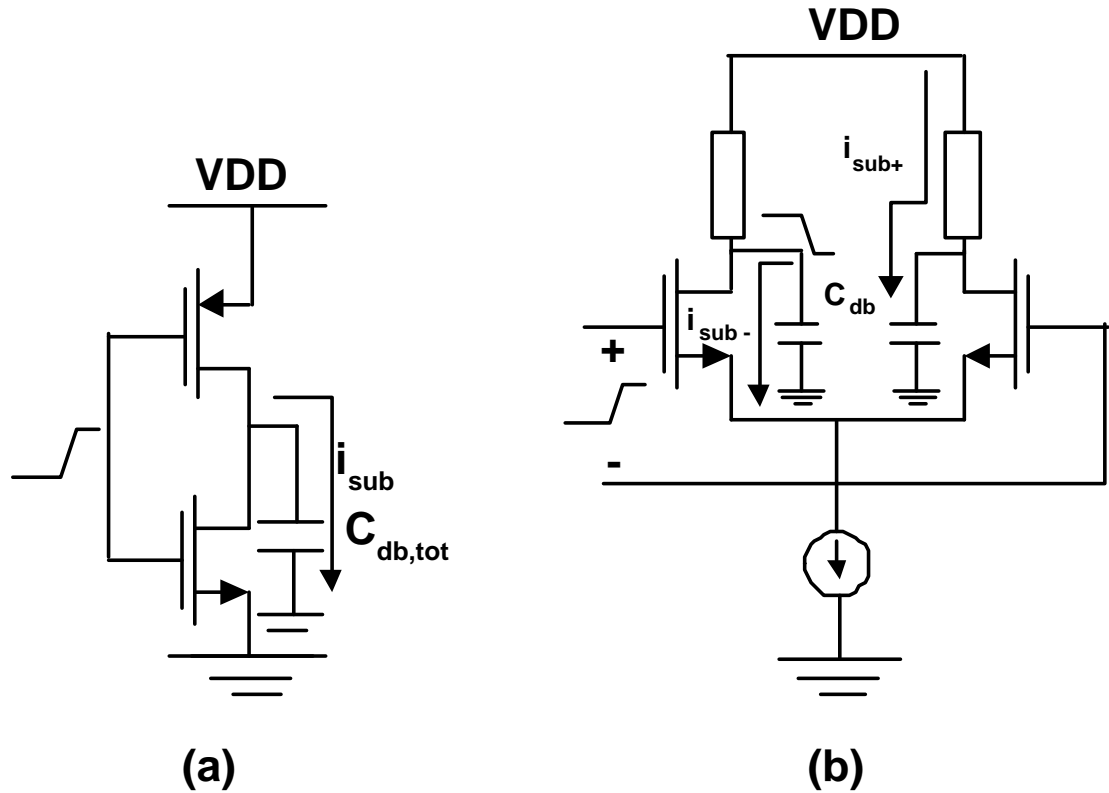
FIGURE 3.5: Substrate current injection in (a) CMOS, (b)CML.

Apart from the above mentioned advantages, differential CML gates give some implementation advantage with the availability of both true and complementary phases of the signal without the need for separate inverters. Finally, their low swing makes them more compatible for low-voltage designs.

3.5.   Pulse-Swallow Feedback Delays

The conventional dual-modulus prescaler with the pulse-swallow architecture is usually limited by the speed of the pulse-swallow operation. In other words, the divide-by-N+1 operation is the speed bottleneck of dual-modulus prescalers. Since the primary

goal of this thesis is to optimize the speed, the feedback loop was analyzed. Referring to the synchronous divide-by-4/5 circuit of Figure 3.2, following the clock edge on which $\overline{Q_2}$ must change , the next valid clock transition needs to accommodate the propagation delay through the gate G and the input stage of $DFF2$. This signal delay can make the divide-by-3 about twice as slow as the divide-by-2 operation.

Some design techniques can be used to reduce these propagation delays in the synchronous division. The combinational gates can be embedded into the first stage of the D flip-flops. Previous implementations of dual-modulus prescalers [24, 14] have incorporated a gate with the flip-flops. The differential current-mode implementation of these "gated" flip-flops has been discussed in the next chapter.

3.6. Ring-Oscillator Speed Analysis

The synchronous portion of the prescaler is the critical design to be optimized for speed. Design optimization of a simple divide-by-two flip-flop begins with reducing the propagation delay of the CML D flip-flops. To estimate the maximum obtainable input frequency that can be divided by the DFF, the toggle flip-flop (divide-by-2) can be regarded similar to a 3-inverter ring oscillator. With the availability of complementary signals, ring oscillators can be made with even number of stages as well. A divide-by-4 circuit is similar in structure to a 4-stage ring oscillator with the complementary output looping back to it's input stage. This parallel between the two circuits is illustrated in Figure 3.6. Theoretically the maximum input frequency toggled by the DFF would be twice the oscillation frequency. However , because of the additional loading of the positive feedback latch in the flip-flops, the input frequencies will be less than twice the oscillation frequencies [24].

The equivalence of ring oscillators and prescalers is significant in analyses for speed-power tradeoffs and in understanding the role of various design parameters such as

(a)

(b)
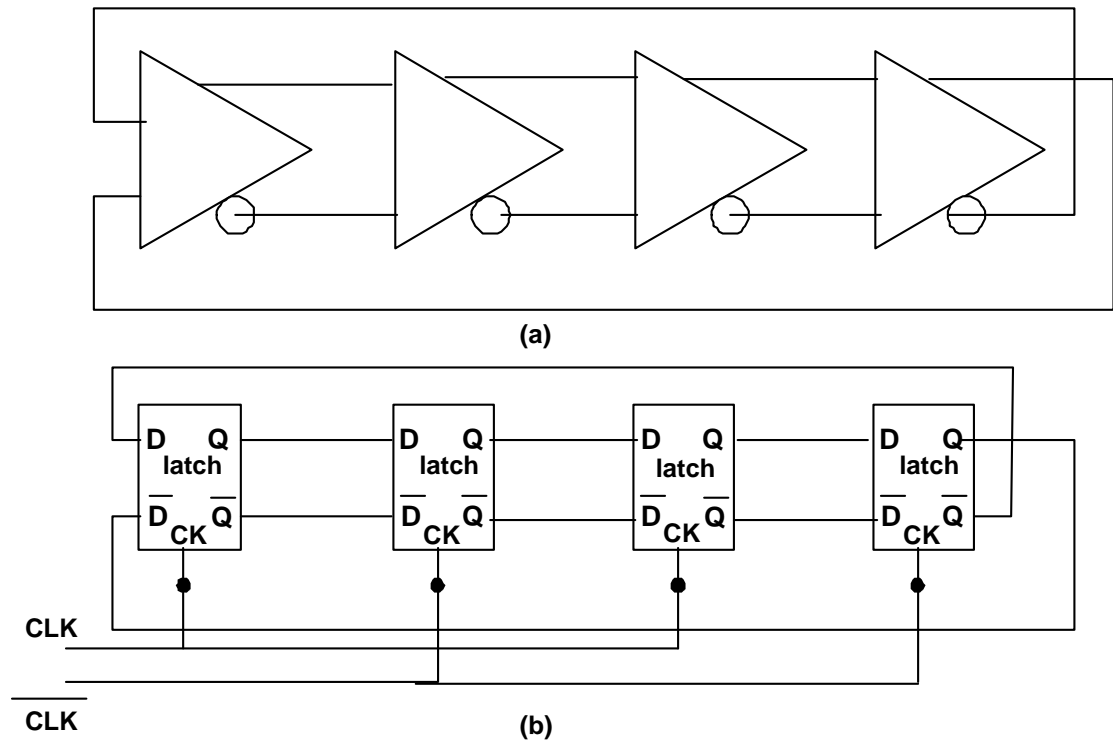
FIGURE 3.6: Analogy between (a) ring oscillator and, (b) frequency divider.

voltage swing, transistor sizes and current consumed in each stage. The results of the analysis are explained in detail in the next chapter.

CHAPTER 4. ANALYSIS, CIRCUIT DESIGN AND IMPLEMENTATION

Having discussed system level considerations in dual-modulus prescalers, this chapter discusses the analysis and transistor-level implementation aspects. The pulse-swallow operation, explained in principle in Section 3.1, is discussed in the context of the divide-by-8/9 prescaler that was designed and implemented.

4.1.   8/9 Dual Modulus Prescaler Operation

The 8/9 dual modulus prescaler is illustrated in Figure 4.1. The synchronous portion, which works at maximum frequency, is the critical block to design. The master-slave D flip-flops FF1 and FF2 perform conventional divide-by-4 in the absence of a "pulse-swallow" signal. Such a control signal can be suppressed by disabling FF3 when Mod-Select signal is inactive. The output of the FF2 is further divided asynchronously to generate a divide-by-8 signal. When this divide-by-8 signal, $Q4$, is combined with the Mod-Select signal appropriately, flip-flop FF3 gets included in the divider feedback loop in such a way that FF1 is forced to hold state for exactly one extra clock period. The output of the synchronous portion now has a duty cycle of $\frac{3}{5}$, i.e., the output Q2 is high for 3 and low for 2 clock periods. $\overline{Q2}$, obviously, follows a high for 2, low for 3 clock-periods trend by virtue of the differential operation of the current mode logic. As the synchronous pulse clocks the asynchronous divider, this translates into Q3 being high for 5, low for 4 pulses (and vice-versa for $\overline{Q3}$). The time period of the prescaler output is now 9 pulses giving it the 8/9 modulus operation.

The pulse-swallow operation is emphasized with a timing diagram (Figure 4.2).
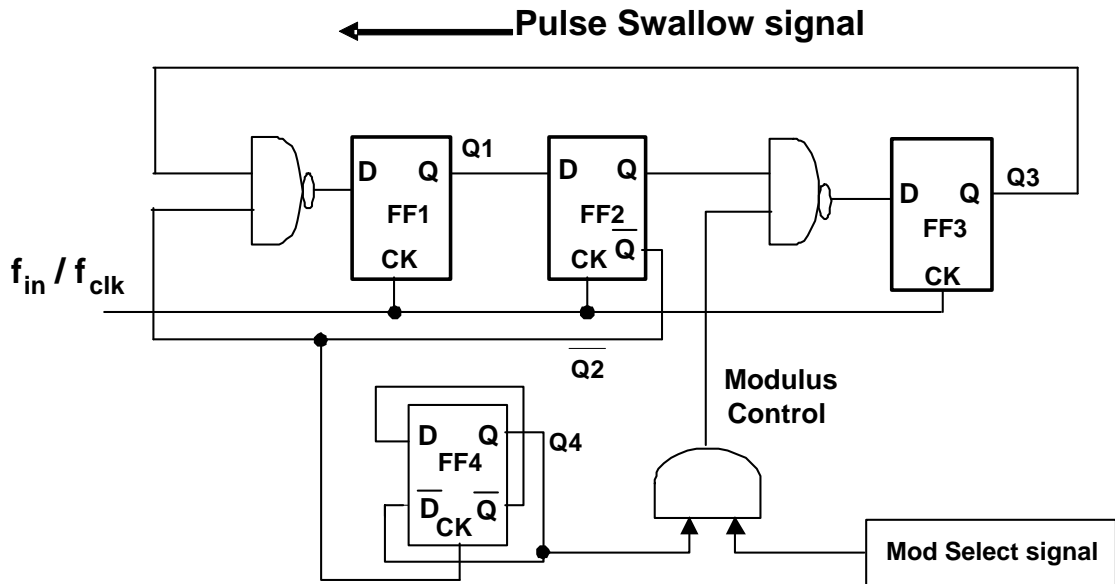
FIGURE 4.1: 8/9 Dual-Modulus Prescaler System.
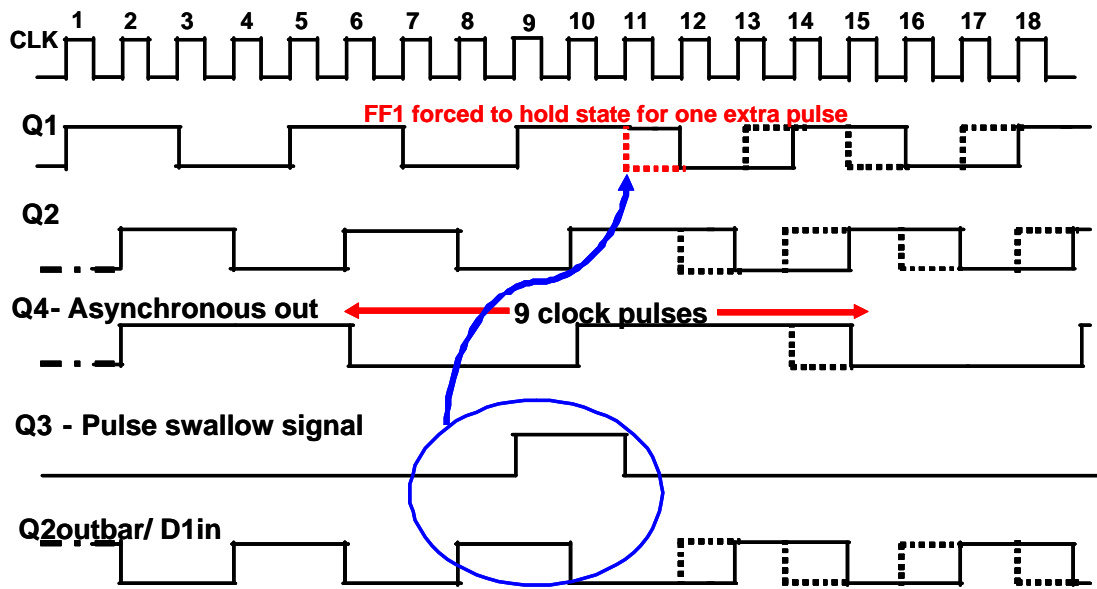
FIGURE 4.2: Timing diagram explanation of pulse-swallow operation.

The prescaler is assumed to be in the ÷5 mode. The output of the flip-flops FF1 and FF2 are, as expected, time shifted by one clock period. The asynchronous divider is clocked by Q2. The asynchronous output, by virtue of an 'AND' operation with Mod-

Select, clocks FF3 so that Q3 is a time-shifted version of $\overline{Q2}$. The pulse-swallow control signal Q3 can thus be considered a 'NOR' operation on asynchronous output and $\overline{Q2}$. The control is used to 'SET' FF1, so that Q1 stays high for one extra pulse more than it would have been without the pulse-swallow operation. The dashed lines show the conventional $\div 4$ waveforms.

4.2.   Design Considerations

The core block to be optimized for speed-power is the synchronous DFF. A detailed analysis of the parameters involved and the optimization has been discussed. As discussed briefly in the earlier chapter, the analysis of the divider structure can be simplified by exploiting the similarity with a ring oscillator. Figure 3.6 explicitly showed this analogy between the two structures. The primary advantage of analysis on the basis of ring oscillator is that the maximum ring oscillation frequency is a clear indication of the speed of the DFFs in the divider. In general, ring oscillators are used to characterize a process because their oscillation frequency depends heavily on the f_T of the transistors. The primary design parameters involved in optimizing the CML Flip-Flops are analyzed in detail in this section.

4.2.1.   Voltage Swing

One of the most significant attributes of current-mode gates over CMOS is its lower output voltage swing. Intuitively, the output node capacitance needs lesser time to charge up and implies faster operation. The formal mathematical equation for the propagation delay of a CML gate may be derived assuming a linear model as shown in Figure 4.3. Assuming symmetry of the differential pair, the initial condition of the

circuit at the beginning of a switching transient initiated by input voltage swing is :

$$V_{o+}(t = 0-) = V_{DD} \qquad (4.1)$$

$$V_{o-}(t = 0-) = V_{DD} - I \cdot R \qquad (4.2)$$
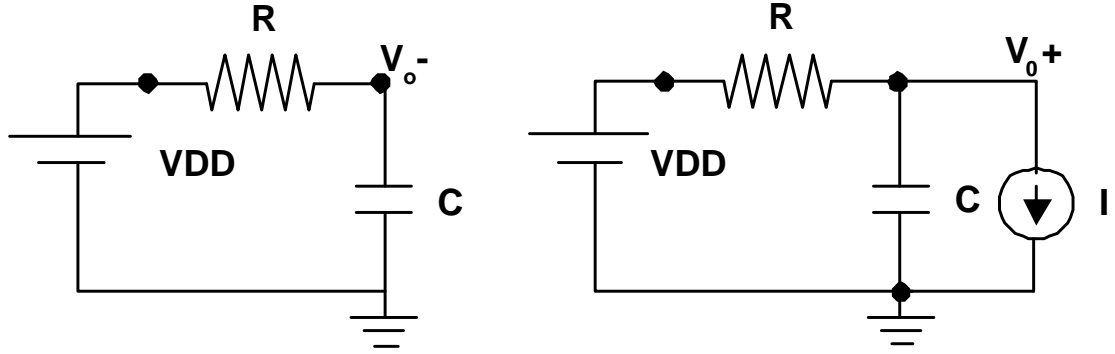
FIGURE 4.3: RC time-constant linear delay model in CML operation.

At the end of the transients, the current is steered from one leg to the other. The output voltages after settling would be

$$V_{o+}(t \to \infty) = V_{DD} - I \cdot R \qquad (4.3)$$

$$V_{o-}(t \to \infty) = V_{DD} \qquad (4.4)$$

Equating transient currents at the output node (assuming instantaneous current switching), we obtain the first order differential equation,

$$C \cdot \frac{dV_o}{dt} + \frac{V_o}{R} = \frac{V_{DD} - I \cdot R}{R} \qquad (4.5)$$

Solving the above differential equation with the initial and final conditions, the output voltage can be expressed as

$$V_{o+} = (V_{DD} - I \cdot R)(1 - e^{-t/RC}) + V_{DD}e^{-t/RC}$$

$$V_{o+} = V_{DD} - I \cdot R(1 - e^{-t/RC}) \qquad (4.6)$$

Propagation delay can be defined as the time taken for the output node to charge/discharge to a desired fraction of the final voltage. For instance, the time taken for the output

to reach within 1% of its final value in the above case of V_{o+}, can be derived from Equation 4.6.

$$
\begin{aligned}
1.01(V_{DD} - I \cdot R) &= V_{DD} - I \cdot R(1 - e^{-t_{99\%}/RC}) \\
\implies t_{99\%} &= R \cdot C \; ln\left(\frac{IR}{0.01(V_{DD} - IR)}\right)
\end{aligned}
\tag{4.7}
$$

The propagation delay evidently depends on the voltage swing $I \cdot R$ in a direct proportion. The voltage swing is a design parameter that depends on other factors as well. In standard CMOS digital circuits, the mid-swing voltage gain is considered representative of the robustness of the circuit to noise [25]. Digital logic requires a point on the DC transfer curve where the gain is greater than 1. This requirement on the gain per stage should be true for a ring of inverters to sustain oscillations as well. The mid-swing gain is given by

$$
\begin{aligned}
A_v &= g_m \cdot R_L \tag{4.8} \\
&= \frac{2I}{\Delta} \cdot \frac{V_{sw}}{I} \\
&= 2\frac{V_{sw}}{\Delta} \tag{4.9}
\end{aligned}
$$

where, V_{sw} is the swing and Δ refers to the over-drive voltage V_{GS} -V_T, or V_{Dsat}.

Another significant reason for higher voltage swings is the response of the positive feedback latch. The output of the preamplifier (differential-pair) of the CML latch, although amplified, still needs to be pulled to the output levels needed to avoid metastability. The latch positive feedback regenerates the output signals to maximum possible swings.

The conventional CML Flip-Flop of Figure 4.4 works similar to a latched comparator with the positive feedback supplementing the gain of the differential pair. The latch-mode time constant in the positive feedback phase has been derived using a linearized model in [26, pp.319-321]. The result derived indicates that the transient response of the latch is represented by the solution

$$
\Delta V = \Delta V_o e^{(A_v - 1)t/\tau}
\tag{4.10}
$$
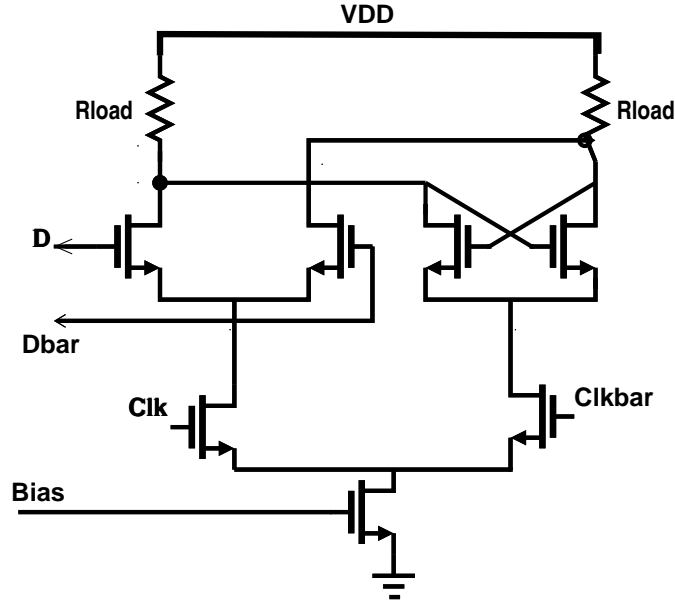
FIGURE 4.4: Current-Mode D-Latch.

where ΔV_o is the initial voltage difference at the beginning of the latch phase.

If it is necessary for a voltage difference of ΔV_{latch} to be obtained in order for the subsequent preamplifier to safely recognize the correct output value, the time required for this to happen can be derived from Equation 4.10 to be

$$T_{latch} = \frac{C_L}{G_m} \ ln\left(\frac{\Delta V_{latch}}{\Delta V_o}\right) \tag{4.11}$$

So if ΔV_o is small, the latch time can be larger than the allowed time to latch (half the clock period) causing metastability. Further, low voltage swings are more susceptible to noise and mismatch. Although not very critical in the case of frequency dividers, this would be relevant for the design of oscillator delay stages.

The upper bound on voltage swings (V_{sw}) is established by biasing conditions of differential pair transistors. When one differential delay stage drives a similar stage, then the differential pair transistor with a high input voltage requires a large enough V_{DS} to

remain in saturation, or

$$V_{DS} \geq V_{GS} - V_T$$

$$V_{DD} - V_{sw} - V_S \geq V_{DD} - V_S - V_T$$

$$\Rightarrow V_{sw} \leq V_T \qquad (4.12)$$

4.2.2. Current Consumption

The current flowing in each stage of the divider/oscillator contributes directly to the static power consumption of the circuit. Since the propagation delay is the time taken for the available current to charge the output node capacitance, the circuit speed is directly dependent on the current through the stage. However, an interesting question that arises is whether there is an upper bound on how fast a circuit can be made to operate if there was unlimited power to burn. In the case of a ring oscillator, if the voltage swing is assumed fixed, scaling up the currents would require (1) reduction in load resistance to maintain swing, and (2) proportional increase in NMOS device sizes so that the over-drive Δ remains same. The increase in the device size implies a proportional increase in parasitic capacitances. Therefore, with a RC time-constant dependent propagation delay, the above two variations nullify the effect of higher current on improvements in speed. At very low device sizes/currents, the gain of the oscillator is not high enough to sustain oscillations.

This result was verified with a simulation on the ring oscillator. Table 4.1 shows that the maximum oscillation frequency of the ring (f_{osc}) does vary with current (I), but not as significantly as one would have liked for the amount of static power traded off. The parasitic capacitances associated with the resistive loads(R) do not scale down proportional to the resistance. So the net RC time constant of the output node starts increasing with current giving diminishing returns in the speed.

TABLE 4.1: Current-Speed relation.

Current I(μA)	W(μm)	load R(KΩ)	f_{osc}(GHz)
50	0.8	12	5.8
100	1.6	6	6.1
200	3.2	3	6
400	6.4	1.5	6.13

4.2.3. Transistor Sizing

The sizes of the transistors in the current mode flip-flops are tightly coupled with the other design parameters of swing and current. The primary considerations involved in deciding the device size are those of speed, voltage swing and current steering ratios.

The RC time constant equations in Section 4.1 suggest that lower device sizes (and hence, lower parasitic capacitance) reduces the delay in each stage. Ideal CML inverters have a perfect current switch that steers current from one leg of the differential pair to the other. In reality, however, some finite current is going to flow in the "OFF" path preventing full current from being available at the output node of the "ON" transistor. Assuming a current of I_{ON} flows through the active transistor and I_{OFF} through the other leg so that $I = I_{ON} + I_{OFF}$, the effective voltage swing is

$$V_{sw} = R[I_{ON} - I_{OFF}] = R[2I_{ON} - I] \tag{4.13}$$

This *current steering ratio* is a parameter that depends on the voltage swing and the device size. It has been observed in [24] that the current steering ratio can be a useful parameter to indicate the robustness of the circuit. The analysis in [24] also accounts for process variations and temperature variations which exacerbate the effect of on CML latches. In this prescaler design too, the device sizes were sized on the basis of a fixed DC current steering ratio (taken as 95%). So the device size involves a tradeoff between

maximum operating speed and robustness to process and temperature variation.

The approach to optimize the design of the inverter stage for each flip-flop began with analysis of ring oscillation frequency with transistor sizes. Simulations results for three particular cases were investigated:

1. Ring oscillator speed with decreasing device dimensions maintaining constant voltage swing and tail current of 100 μA (Figure 4.5(a)). As the percentage current steering will be lower for smaller transistor sizes, the load resistance is higher than $\frac{V_{sw}}{I}$. The increases in resistance is, however, not significant enough to setback the improvement in node capacitance. So, although the slope of the frequency variation flattens out at lower device dimensions, the general trend encourages smaller size for faster operation.

2. Ring oscillator speed with decreasing device dimensions maintaining current steering percentage fixed and the total current at 100 μA constant (Figure 4.5(b)). The load resistance used is assumed ideal for simplicity. The above discussion indicates that the current steering requirement imposes need for higher device dimensions, and so, the maximum oscillation frequencies are lower. The trend remains the same as in the case of the fixed swing

3. The above simulations, when run with the real resistance RBH2 available in the National BiCMOS8i process, shows a "sweet-spot" at very low device sizes (Figure 4.5(c)). This inflection occurs because at very low device sizes, the load resistance needed to compensate for the lower steering is so high that the parasitic capacitances associated with these resistances negate the reduction of the transistor parasitic capacitances. The parasitic capacitances associated with the load resistance are small and may be swamped out when interconnect capacitances are included in a real simulation with extracted netlists.

4. The fourth and most relevant set of simulations was done with the divide-by-4 circuits. The input clock frequency was stepped up for different device sizes in the

flip-flop until correct division operation observed. Although the ring-oscillator is similar to the divider circuit, there are some marked distinctions. The additional positive feedback latch stage in the flip-flop not only loads the output of the CML preamplifier stage, but also helps the speed with its gain. The differential pair and the latch have similar transistor sizes for optimum current steering. The design had to account for process and temperature variations as well. The schematic of the final current-mode D flip-flop of the prescaler, designed on the basis of above considerations is as shown in Figure 4.6. The clock and signal swings were set to about 0.6 V with current through each CML latch at 100 μA.
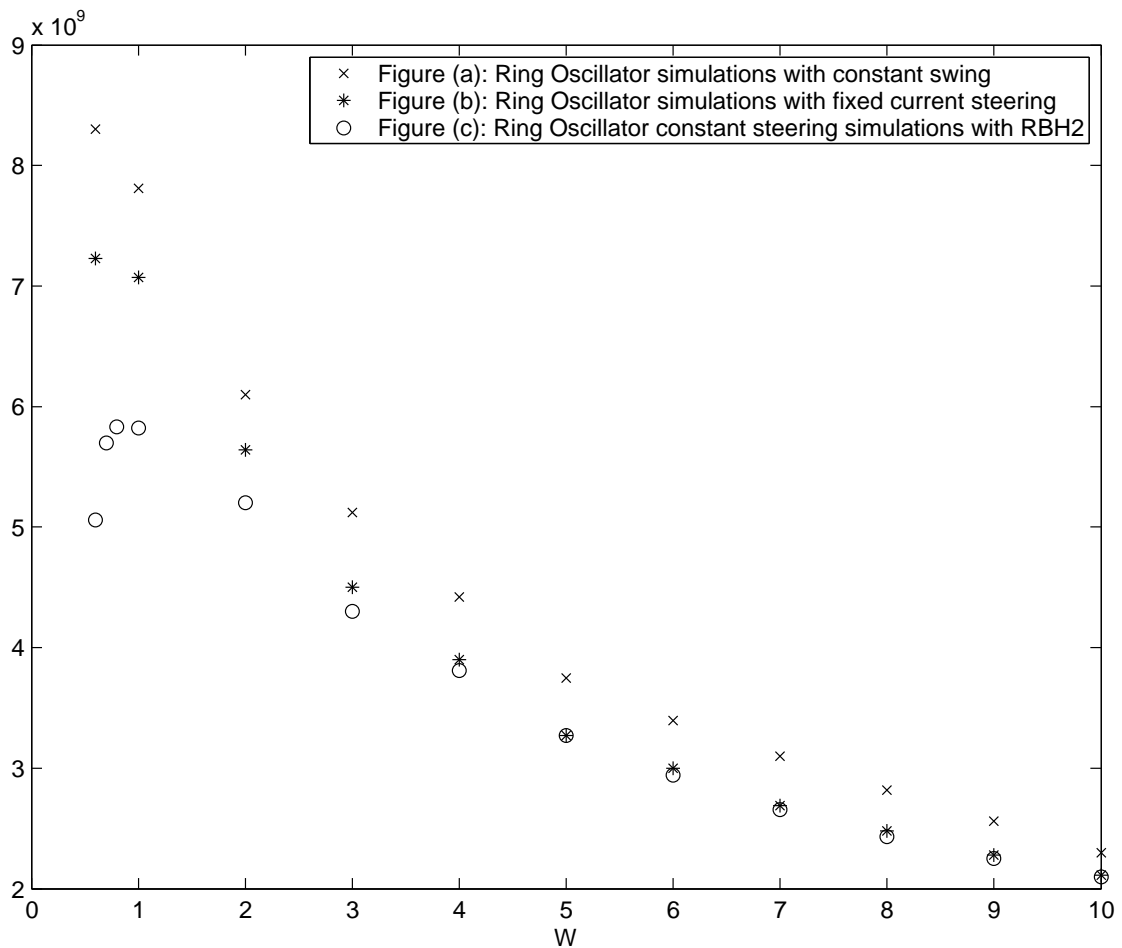
FIGURE 4.5: Ring oscillator simulation comparisons.
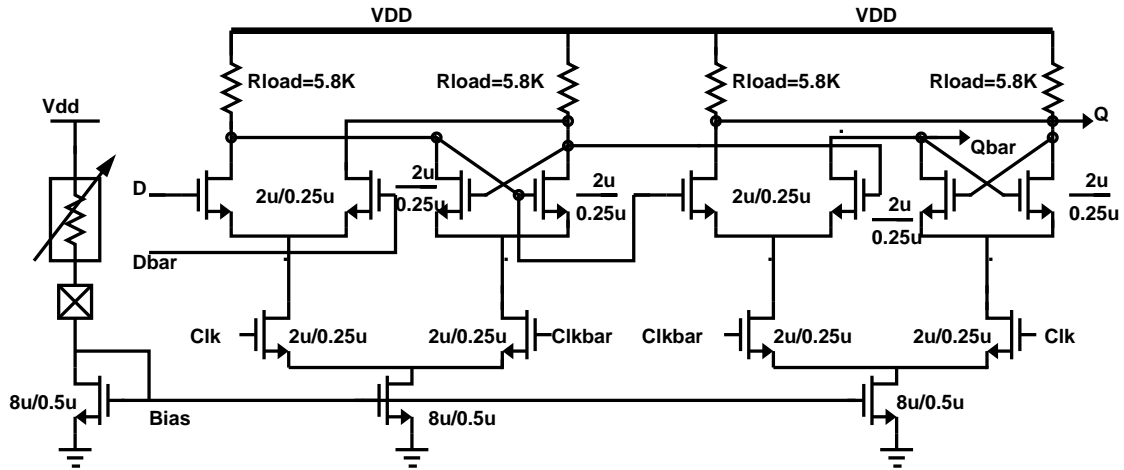
FIGURE 4.6: Optimized D flip-flop

4.3. Implementation Of Pulse-Swallow Logic

As mentioned earlier, the speed bottleneck of the dual-modulus prescalers is in the divide-by-N+1 implementation. This is obvious considering the fact that the N+1 modulus division requires the divide-by-N signal (and hence, the delays associated with it) as well as the delay in generating the pulse-swallow signal. There have been some clever design techniques to reduce the pulse-swallow delays [14, 27, 15] using bipolar ECL and ECL-like differential logic. Merging the logic gates into the flip-flop saves power and increases operating speed. However, some of the above methods, especially the previous generation prescaler implemented in the National BiCMOS7 process and its MOS current-mode equivalent in [15] have their disadvantages.

The gated D-type master-slave CML latch is shown in Figure 4.7(a). The *reset* signal needs to be combined with the divider signal to pull the ouput node to a logic 'low' state. Unlike the simple DFF where the signals are differential and symmetrical, the OR function of these gated flip-flops requires that the input signals compare their levels with a reference voltage to determine whether whether the signal is high or low.

In current-mode logic, the signal swing is low and the DC value of this reference may tend to shift around due to process variations. The way the reset operation works is based on providing a dominant pull-down path through the reset transistor in parallel with the DFF signal transistor. The disadvantage of such a technique is that since the reset operation is essentially single-ended and asymmetric, we loose many of the common-mode noise immunity advantages discussed in the previous chapter. Also, since it requires a dominant pull-down 'reset' transistor, this needs to be 4× or 5× wider than the conventional differential pair/latch transistors. This larger device for the one extra pulse out of the N (divide modulus) pulses loads the differential pair and slows down the prescaler operation. Any logic that requires a "fight" between two signal paths cannot be robust.
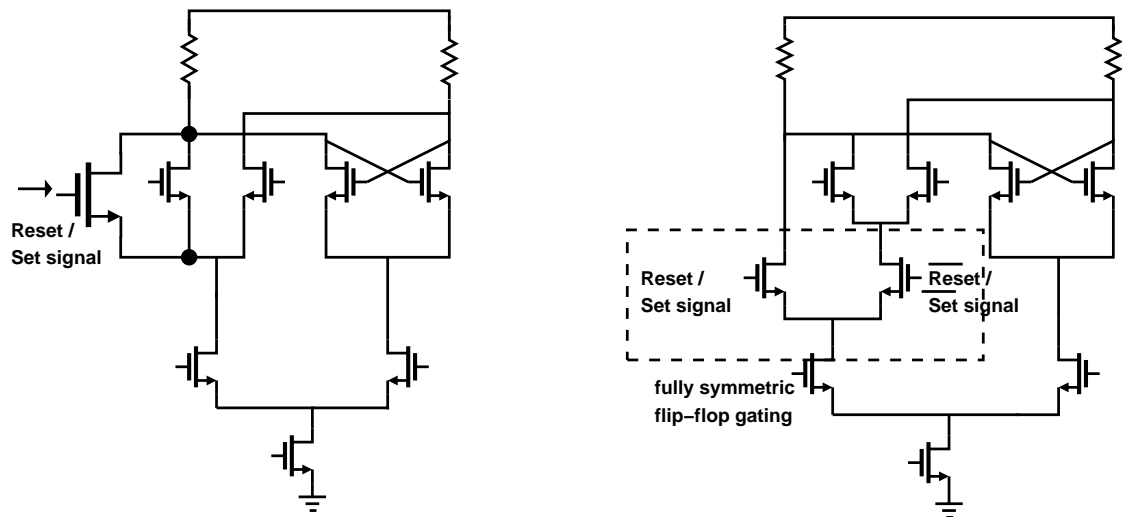
FIGURE 4.7: (a) Parallel pull-down latch structure for gated flip-flop.(b) Fully symmetric flip-flop gating.

A more symmetric implementation of the reset/set operations on the gated flip-flops is shown in Figure 4.7. The principle of operation is based on the idea of stacking CML gates (their low swing allows multiple stacking) to save power [14]. Since a CML gate has only one structure, and all logic operations can be derived from one basic CML cell, the reset/set operation is defined by combinational logic. For example, the

conventional 'reset' operation is an 'AND' of signal with 'LOW' and 'set' operation is an 'OR' with 'HIGH'. A CML gate's structure is inherently asymmetric with respect to the output node loading. Careful layout by appropriate source/drain sharing of transistors and use of dummy devices can alleviate problems due to this load mismatch.

4.4.   Asynchronous Flip-Flop

The asynchronous divide-by-two flip-flop is simpler to design as it works at $f_{clk}/4$. To design the asynchronous DFF4, the device sizes of DFF1 are retained, but the current is scaled down. Since the output of DFF4 needs to be combined with the modulus-select signal to control the pulse-swallow flip-flop DFF3, the output swing should be the same as the signal swing in the rest of the circuit. The load resistance needs to be scaled up to compensate for the lower current in the stage.

Figure 4.8 shows the asynchronous $\div 2$ stage schematic. This stage also needs an input buffer because the divide-by-4/5 signal that clocks this stage is usually not a clean reference. Simple CML inverter buffers are used with source-followers for level shifting the signal level to the clock bias levels. In this design, the targeted operation was 8/9 dual-modulus and so, requires only one such asynchronous divider with the 4/5 synchronous stage. Higher moduli can be obtained by cascading more toggle flip-flops (TFFs).

4.5.   RF Buffer

The design of the dual-modulus prescaler was implemented entirely in CMOS, which was one of the key objectives of this work. The prescaler in the PLL loops are driven by the VCO output. It is typical to buffer the VCO outputs to shield the noise-sensitive block from the digital switching in the prescaler. The specifications on the
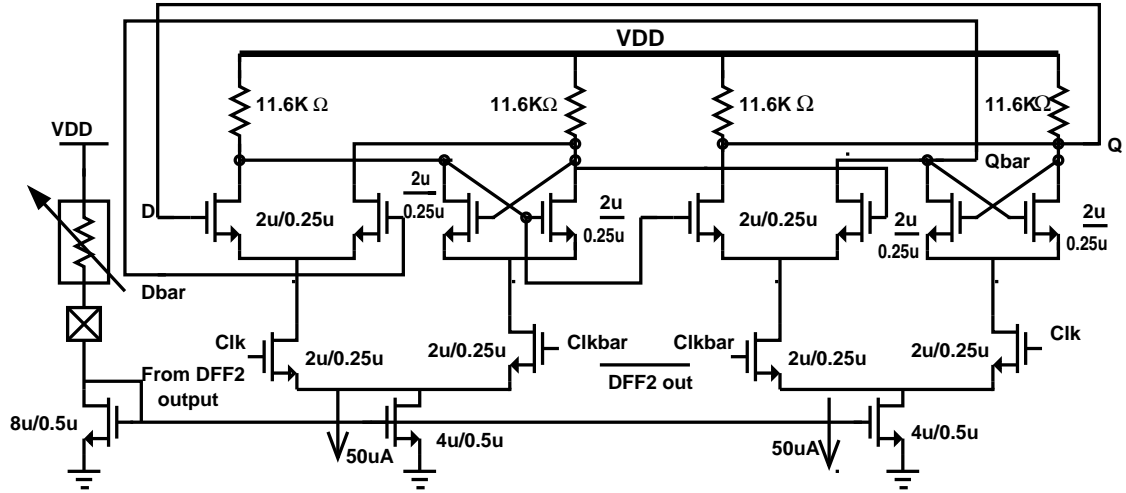
FIGURE 4.8: Asynchronous divider.

dual-modulus prescaler requires operation for signals in the range -20 dBm to 0 dBm power level. This needs to be converted to $\approx$ 0.6 V single-ended peak-to-peak swings to clock the prescaler. The above requirement imposes a gain specification of about 26 dB ($\times$20) at the targeted signal frequencies of 2.5-3 GHz, which translates to gain-bandwidth product requirements to the order of 50 GHz. The buffer was implemented using both CMOS and bipolar technology and the performances compared.

4.5.1. CMOS RF Buffer

The gain requirement for low-power input signals implies the need to cascade multiple gain stages. The primary requirements on the output clock signal generated by the RF buffer are:

- The output signal should have a high rise time so that the the the differential pair transistors of the prescaler spend lesser time in the linear region. This achieves lower total noise contributed by the transistors [5]. The output waveform should effectively be as close to a square waveform as possible with signal swings of 0.6 V.

- Based on DC biasing conditions, if the clock signal amplitude is large, the differential pair transistors driven by the clock could swing into triode region. This is undesirable for robust current switching.

    With the above requirements in consideration, a cascade of amplifiers was designed. The first stage does a single-ended to differential conversion and the remaining stages provide increasing gain (Figure 4.9). In order to limit the output swing, each of the gain stages has a common mode load resistance with about 0.9 V drop across it. This is effectively like working with a 1.6 V supply. To shape sinusoidal waveforms into square waves, more stages of amplifiers and higher current needs to be burned to sharpen the edges of the clock.
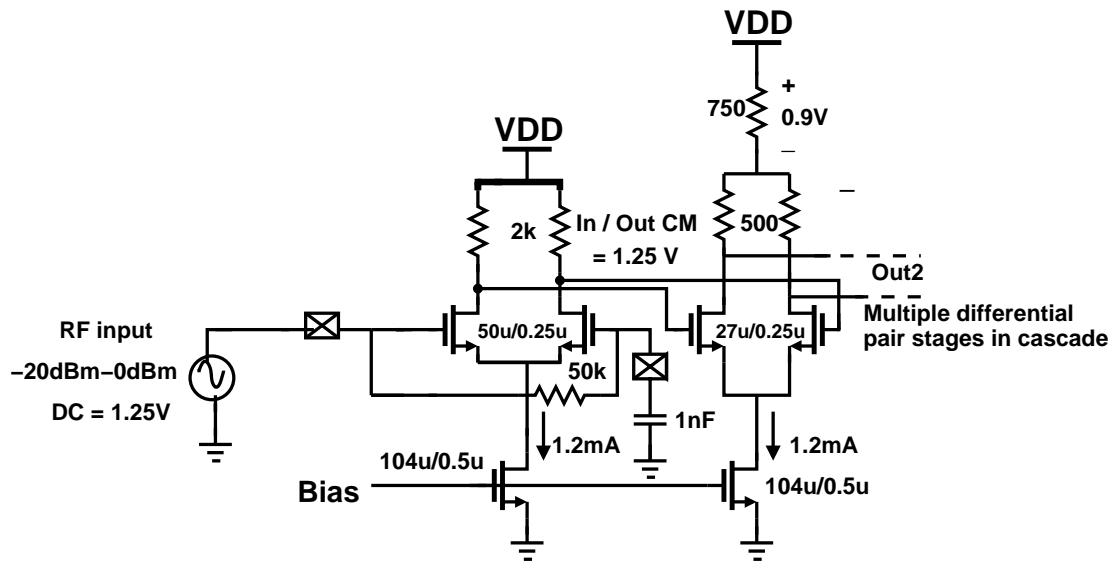

FIGURE 4.9: CMOS RF buffer.

    The load resistances were chosen to allow no more than 0.6 V swing across them. The device sizes are chosen using parametric sweep of the transistor widths to obtain a high gain, square waveform for a fixed current and swing. If the transistor W/L is large, its g_m is high, but it loads the previous stage as well. So the parametric sweep yields an

optimum device size for high gain. Only the first differential to single ended stage does not require the supply level-shifting resistance.

In order to generate square waveforms for worst case conditions (slow transistors, hot temperature conditions), 12 gain stages were required and with about 1.2mA in each stage. To characterize the "square"-ness of the waveforms, the rise/fall time of the current through the final output stage (time to swing from 5% to 95% of final value) as a proportion of the time period (which needs to be ≥ 2.5 GHz). In the worst corner of slow process/hot temperature, the rise time with the former criterion is about 26% of the time period. Effectively, the differential pair is in the linear region about 52% of the time. For nominal case operation, this rise time is 15% of time period.

The output swing can achieve the desired 600 mV with 6 gain stages. The additional stages and larger current can be traded-off with the noise criterion. Some of the clock buffer output waveforms (Figure 4.10) are shown here to illustrate the performance of the design. This buffer's power consumption is unacceptably high and the number of cascaded stages would increase the clock jitter.
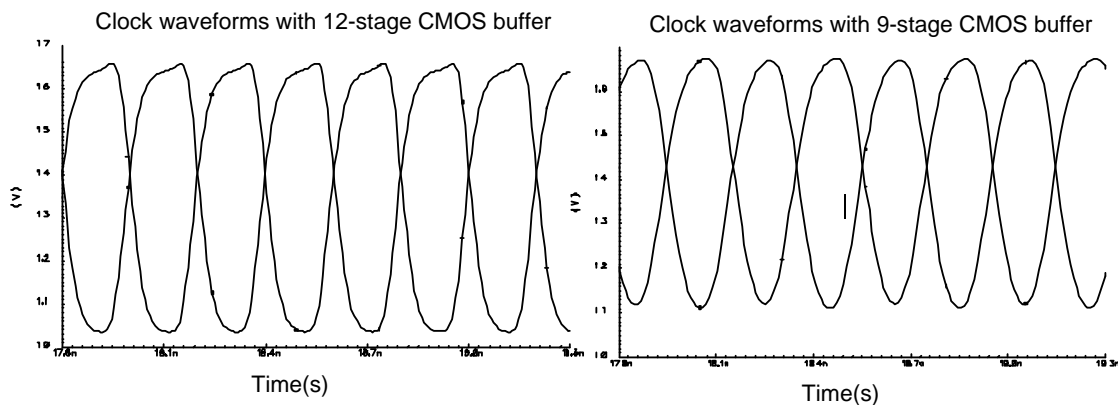

FIGURE 4.10: Clock waveforms : CMOS buffer output.

4.5.2.   BiCMOS RF Buffer

The significant advantages of bipolar transistors for high frequency analog applications were outlined in the previous chapter. The much higher $\frac{g_m}{I_d}$ of the device makes it possible to obtain higher bandwidths. The RF amplifier/buffer shown in Figure 4.11, can be implemented with just two gain stages using the available NPN devices for the differential pairs. The primary reason why the power supply had to be level shifted in the CMOS implementation is that source-follower level shifters have an AC gain at 2.5 GHz of about 0.4, therefore degrading the gain obtained from the differential amplifiers. In the case of bipolar transistors, the small-signal AC response of emitter followers has a wide bandwidth so that the small-signal gain is approximately unity. This enables full utilization of the power supply.
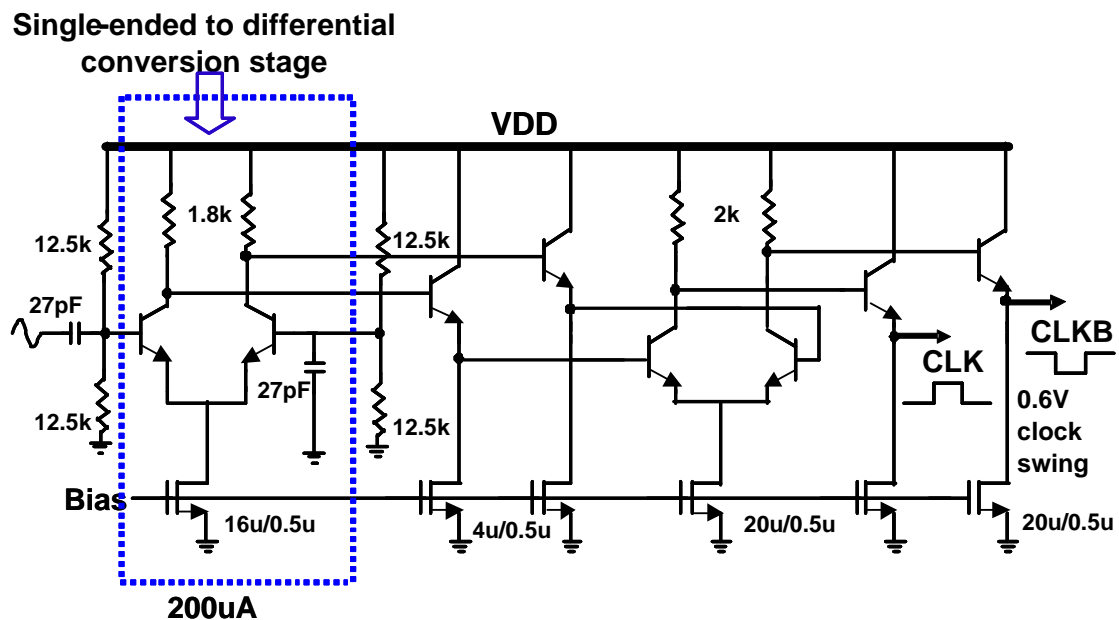

FIGURE 4.11: BiCMOS RF Buffer.

The design issues for bipolar buffer are listed below :

• The bias currents should be set such that the transistors work to the left of the

peak f_T point. The f_T vs current characteristic was obtained for the available npn transistors and the bias point point chosen accordingly.

- The swing of the clock signals must be maintained such that the collector-base junction is not forward biased at any instant. The resistances and the bias currents must be set such that the largest input signal amplitude does not saturate the npn transistors.

- The frequency response of the bipolar buffer must allow for sufficient gain to switch currents completely at the final output stage. The total capacitance at the common emitter junction of the differential pairs determines the high frequency response of the and should be minimized. As bipolar transistors suffer from large collector-substrate capacitance, MOS current sources are used to bias the transistors.

The single-ended to differential conversion cannot be implemented as in the case of the CMOS buffer due to the mismatch introduced by base currents. So the common mode of the differential pair transistors of the first stage are set by independent references, generated by resistive voltage division.

4.6. Output Buffer

The dual-modulus prescaler in a conventional frequency synthesizer drives a counter or the phase-frequency detector of the PLL. So the load of the prescaler is usually capacitive, i.e., the gate capacitance of a few transistors. For this test chip, since the dual-modulus prescaler was implemented as a stand-alone block, it needs larger drive capability. The output buffer needs to be designed to drive a 50 Ω resistance,and at least a 5 pF capacitive load. The biggest concern with driving these huge loads is the amount of current that is drawn during voltage swings of the prescalers output.

To match the output impedance as well as to supply this large current a source-follower/emitter-follower buffer may be used. The currents may be of the order of a few

milliamps. MOS implementations using source-followers required large device sizes to support the current. These devices may then load the prescaler, affecting its operating speed. A more feasible alternative to satisfy this high drive capability is to use bipolar transistors (Figure 4.12). To avoid large DC currents due to the output resistance, the buffer is AC coupled to the load. The sourcing current to charge the output node needs to be supported by the transistor of the emitter follower. Multiple transistors are used in parallel so that each transistor only carries current corresponding to its optimum f_T. The sink current, however, is provided externally through resistances.
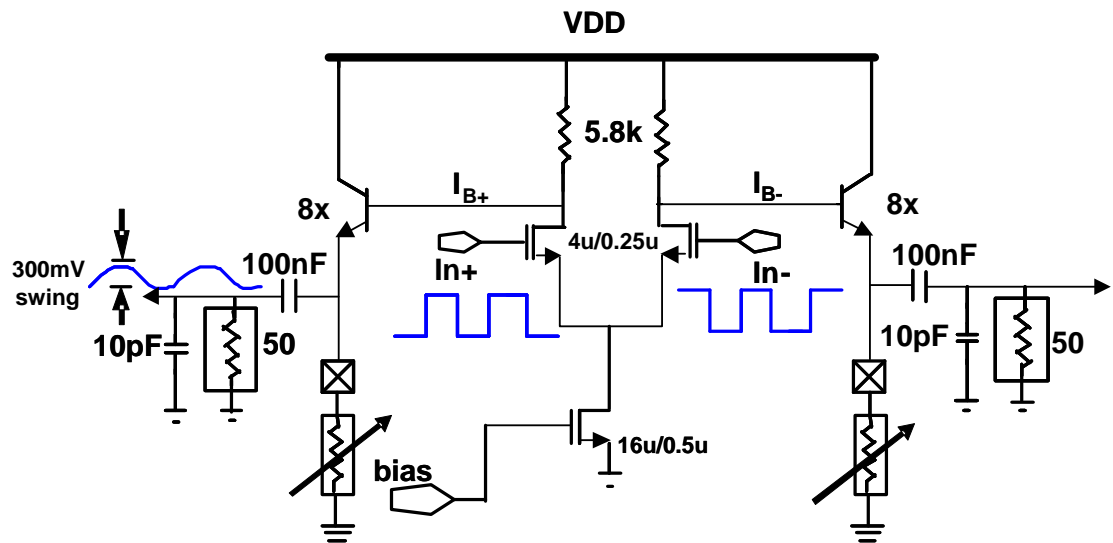

FIGURE 4.12: BiCMOS Output Buffer

A new issue that crops up with the use of bipolar drivers is the large base current component. With a collector current gain β of 100, the base current could be as high as 250 μA. The isolation buffer between the prescaler from the bipolar devices needs to supply the extra base current into the input of the emitter follower.

4.7.   Layout Considerations

The layout of RF building blocks is very important in realizing expected performance. The main layout considerations and techniques to optimize the dual-modulus prescaler system have been summarized below.

4.7.1.   Symmetry Considerations

One of the highlights of implementing the dual-modulus prescaler system with current-mode logic is the advantage of common-mode noise immunity achieved by its differential and symmetric operation. Effect of random mismatches could be significant in introducing clock jitter, especially with small device sizes. Since the prescaler and every flip-flop is desired to have large-signal symmetry, the dummy transistors of Figure 4.13 ensure the output nodes of the differential amplifier in the first level stacking see the same signal swing. The obvious drawback is that the capacitive loading is now asymmetrical on the output nodes. However, since the transistor sizes are quite small, the dominant capacitances of these asymmetric output nodes is the gate capacitance of the positive feedback latch and the interconnect capacitance. So this asymmetry is expected to have a lesser significance compared to signal path symmetry. Conventional techniques, such as sharing common source terminal of differential-pairs and implementing all transistors with respect to unit transistors, have also been used in drawing the layout.

4.7.2.   Synchronous Divider Floorplan

As explained previously, the pulse-swallow feedback limits the speed of operation. So, to optimize the layout with speed consideration, the delays in the critical synchronous divider path need to be reduced. This circuit is dominated by interconnect

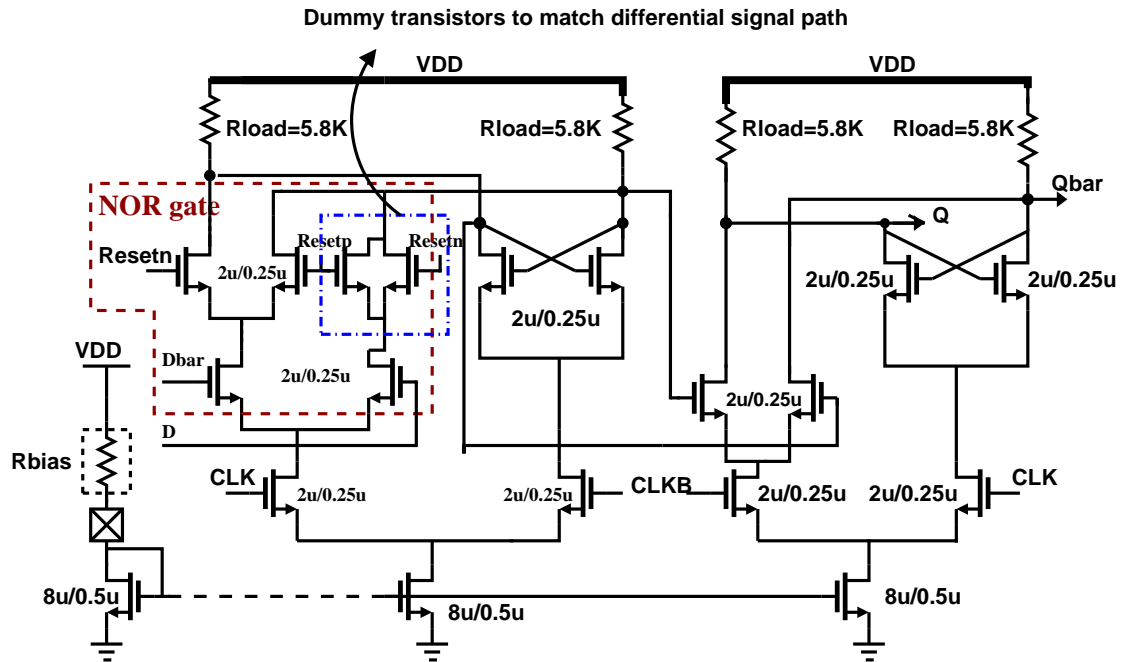**Dummy transistors to match differential signal path**

FIGURE 4.13: Flip-Flop3 with dummy devices to maintain signal symmetry.

capacitance, so it is imperative to ensure that wiring delays are minimized, especially in the synchronous divider. It was obtained from analysis of the critical delay paths that placing the pulse-swallow flip-flop DFF3 as close as possible to the first stage of the flip-flop DFF1 reduces the feedback delay. The synchronous divide-by-4 operation between DFF1 and DFF2 will have a lesser delay compared to the propagation path of the divide-by-5. A simple floor-plan of the layout of the dual-modulus prescaler, not including the modulus-selection logic, is shown in Figure 4.14.
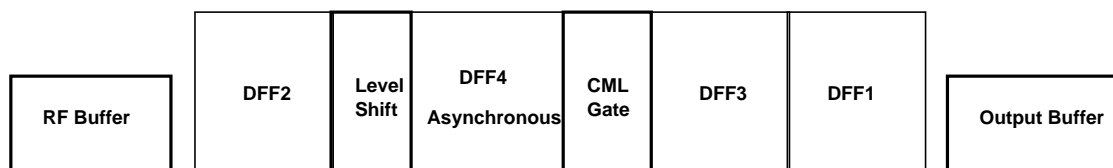

FIGURE 4.14: Floor plan to optimize pulse-swallow feedback delays.

It was verified from simulations on the extracted netlists that this placement of the flip-flops and combinational logic ensures breakdown of division operation for both mod-

ulus at approximately similar frequencies. If the layout floor plan had been implemented as in the sequence of the block diagram of Figure 4.1, the divide-by-9 operation breaks down approximately 300 MHz input clock frequencies below the divide-by-8 modulus division. It is obvious that the higher operation of the divide-by-N stage alone is not desired. By shuffling the arrangement of the flip-flops, we are trading off the speed of the synchronous division with the speed of the pulse-swallow division.

4.7.3. Minimization of Interconnect Capacitance

As the interconnect capacitance at the output node of each differential amplifier/latch dominates the RC propagation delay, a study of the metal layer used to route was performed. Further, we have a choice between two types of routing–using a top level metal (M4, M5) that is routed above the rest of the layout to connect end-to-end flip-flops, or to route over a longer path using lower-level (M1, M2,or M3) metal avoiding signal crossing. It was observed that routing using M5 right over the the rest of the layout has lesser RC parasitics. Parasitic coupling between the signal lines and the power buses can be avoided with this strategy as well. Simulations using extracted netlist confirm the above deductions. Another important rule that was followed in the layout was to isolate the clock lines from overlapping signal lines to the maximum extent possible to avoid disturbing the critical reference frequency.

The top-level layout of the dual-modulus prescaler is shown in Figure 4.15. The prescaler, input and output buffer layout is captured in Figure 4.16.
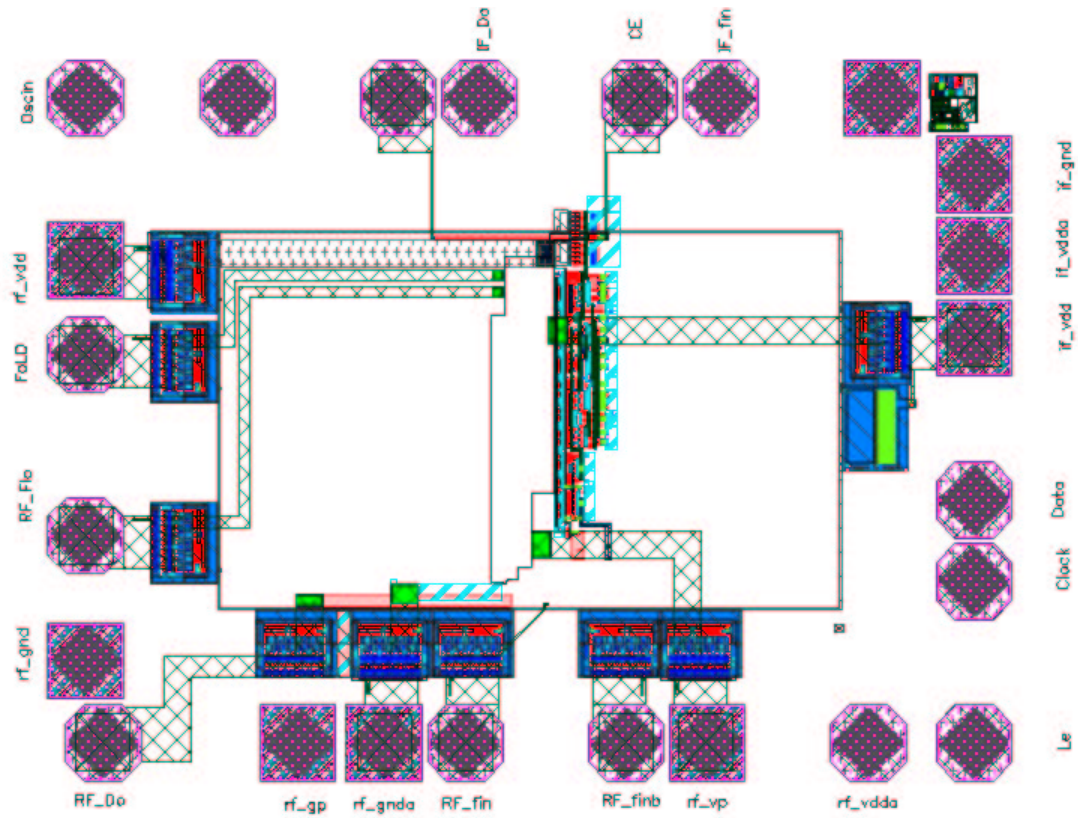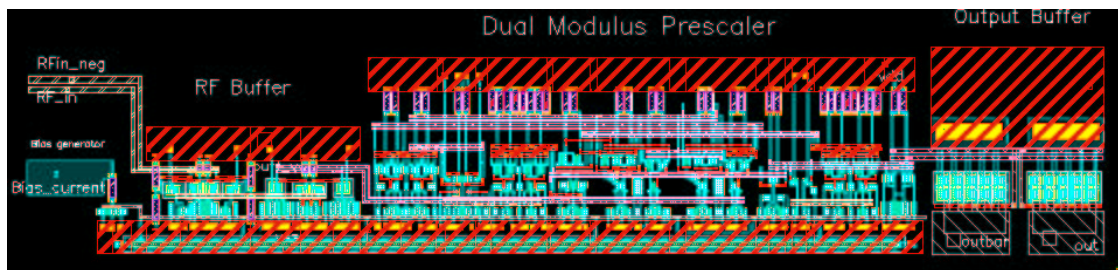
FIGURE 4.15: Top-level chip snap-shot.

FIGURE 4.16: Layout.

CHAPTER 5. SIMULATION AND MEASUREMENT RESULTS

This chapter is a discussion of the simulation results obtained and a performance evaluation of the prototype chip. The dual-modulus prescaler, analyzed and described in the previous chapter, was designed and laid out using the CADENCE simulation environment. A chip was fabricated using the National BiCMOS8i 0.25 μm, 5-metal layer process. The chip was mounted on a 20-pin ultra-thin chip scale package (UTCSP) provided by National Semiconductor. The die photograph of the fabricated chip is shown in Figure 5.1.
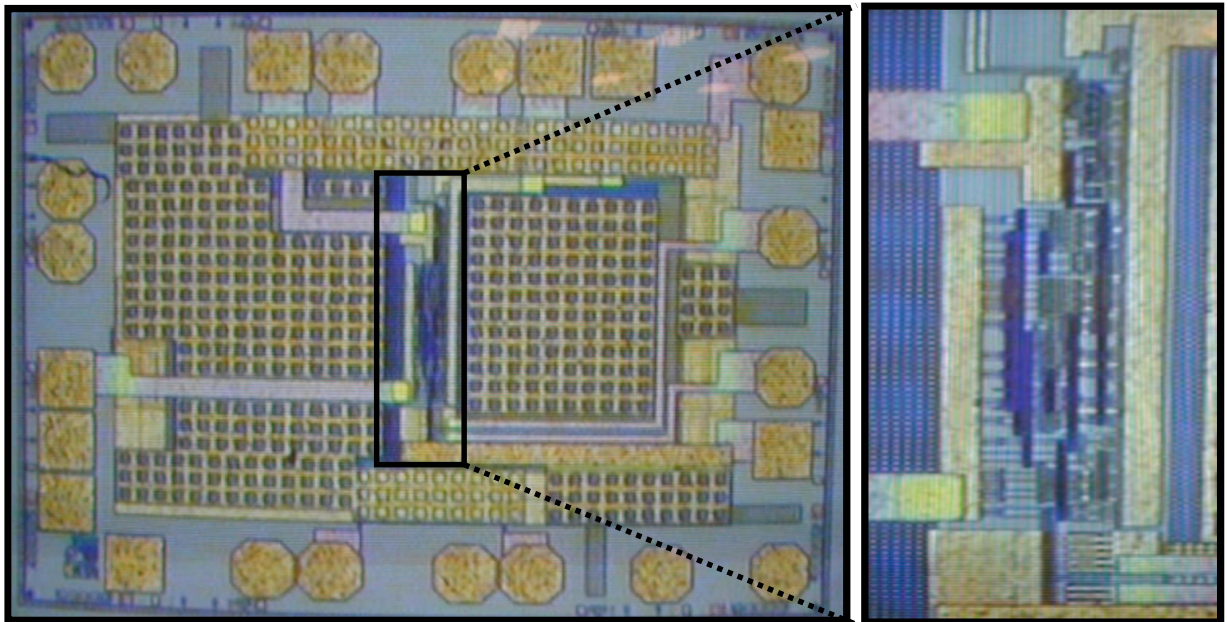

FIGURE 5.1: Die photograph of OSU-Prescaler test chip.

5.1. SpectreS Simulations

The overall system that was simulated is shown in Figure 5.2. The dual-modulus prescaler and the input/output buffer blocks have been discussed in the earlier chapter. The modulus-selection logic is usually determined by the program counter and swallow counter data in integer-N PLLs or, by means of a randomized modulator output, as in the case of a $\Delta\Sigma$ fractional-N synthesizer. For the prototype testing a worst-case switching was decided upon. Worst case switching happens when the modulus is changed between N and N+1 every other complete cycle. This can be easily realized using a divide-by-2 asynchronous block clocked by the 8/9 modulus output from the prescaler.
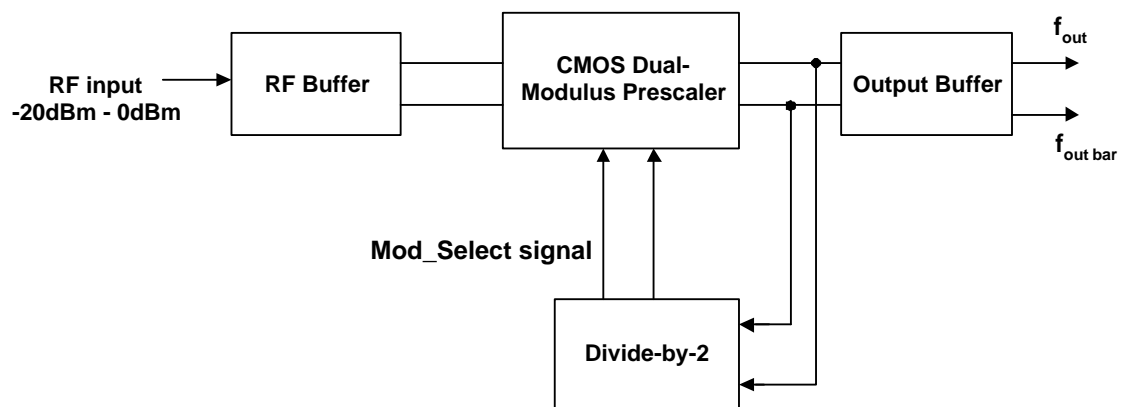
FIGURE 5.2: Top-level of dual-modulus prescaler implementation.

The operation of the prescaler was verified by looking at the transient waveforms and checking for dual-modulus operation. The slowest case operation was obtained with the slow process transistor models and hot temperature (85^0C). A typical transient simulation result of the prescaler operating at 2.75 GHz is shown in Figure 5.3.

Although the fast-case transistor models and higher input power levels appear better for speed performance, a possible issue could be the robustness of the buffers. As shown in the transient waveform of Figure 5.4, the bipolar buffers tend to saturate at
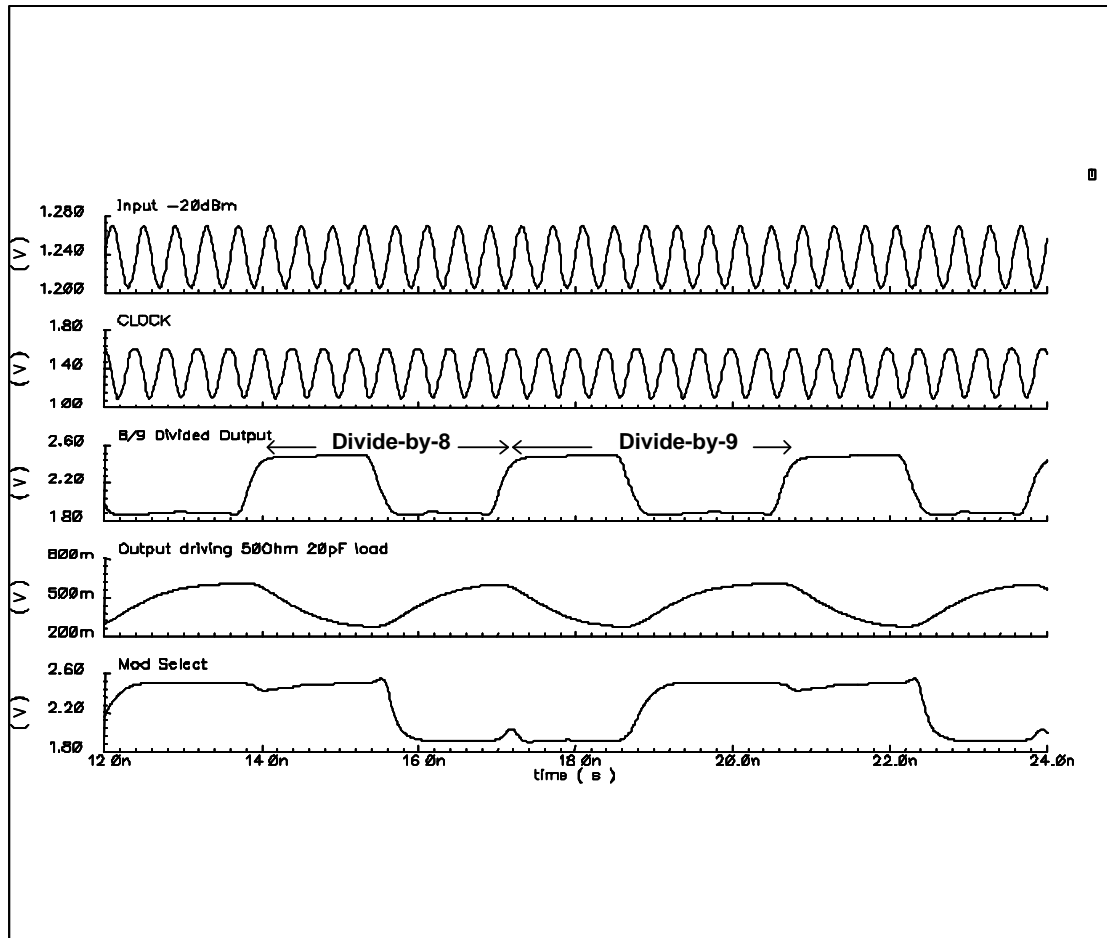
FIGURE 5.3: Prescaler output waveform for slow process/hot temperature corner.

0dBm input signal power levels.

It was observed in simulations that the divide-by-9 limits the operating frequency of the prescaler. Investigations by following the waveform delays through the pulse-swallow feedback loop indicate that the propagation delays through the asynchronous divider output and modulus-selection logic causes a time-delay between the output of DFF2 and the RESET signal on DFF3. This delay suppresses the pulse-swallow signal even when mod-select has been asserted. The simulation results indicating measured delays encountered in the pulse-swallow operation is shown in Figure 5.5.
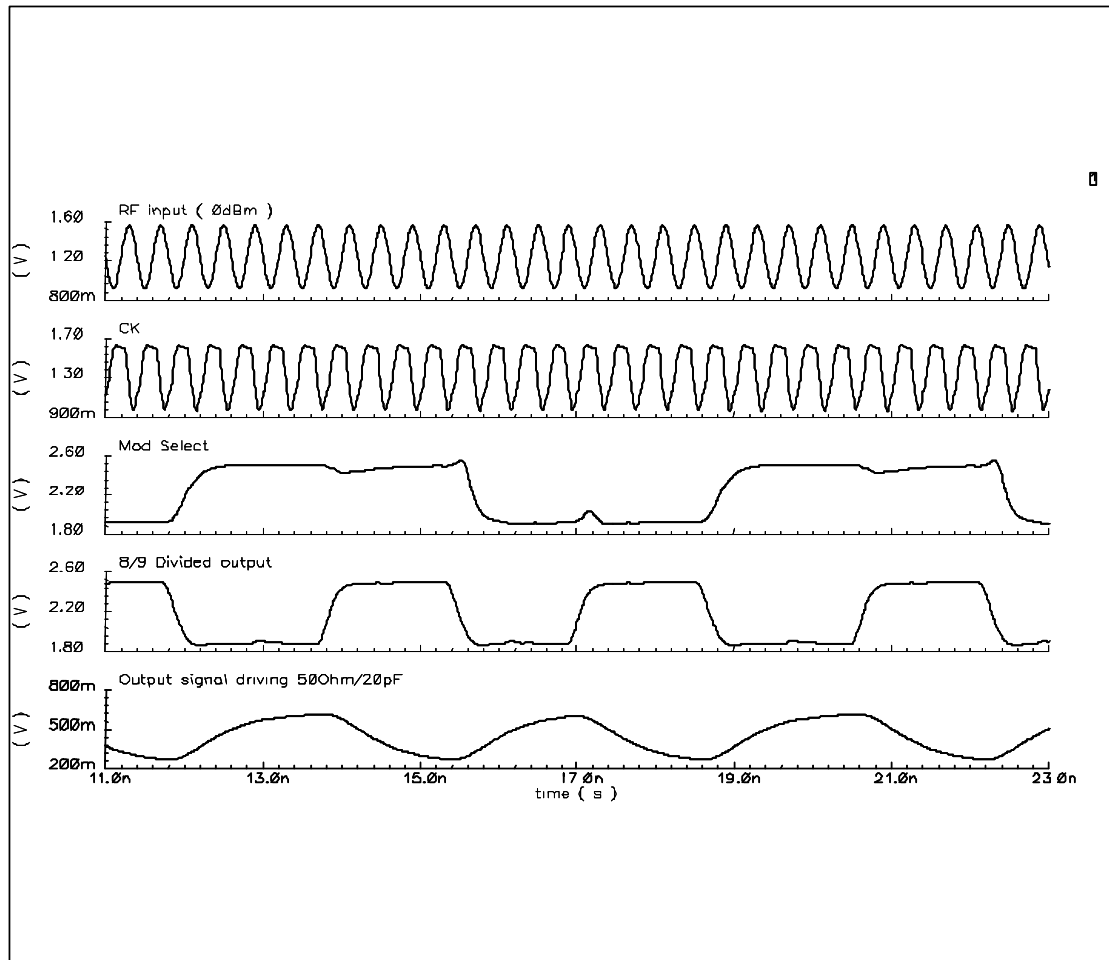
FIGURE 5.4: Prescaler output waveforms for 0 dBm input, fast/cold operating corner.

5.2. Measurement Set-Up

The prescaler measurement set-up is shown in Figure 5.6. The chip's input/output pins have been labelled for easy identification. The input frequency to the chip is through a signal generator to obtain progressively higher clock frequencies by ramping the control voltage. The power levels are also to be controlled to verify operation of the RF buffer. The control to vary the speed of the prescaler is by varying the current to each stage of the synchronous divider. The current bias is decided by external reference voltage or resistance. Since the bias currents of the RF buffer, prescaler and output buffer are

Total propagation delay ~330ps

FIGURE 5.5: Dual-modulus division operation breakdown.

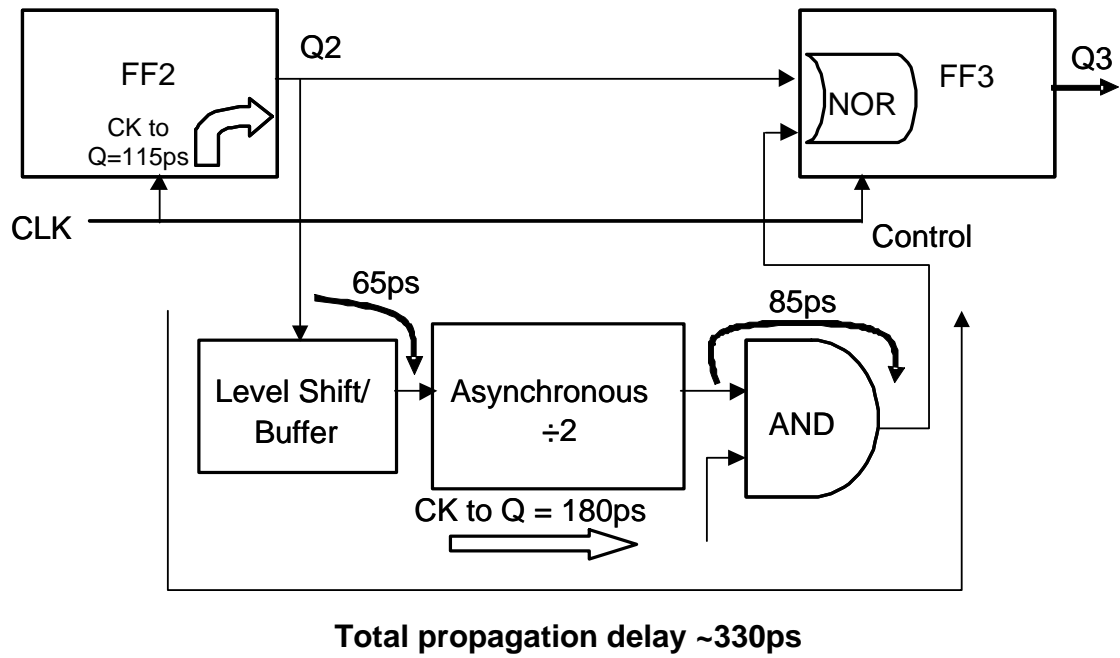all mirrored with the same bias transistor, increasing the bias current worsens power consumption of all three blocks.
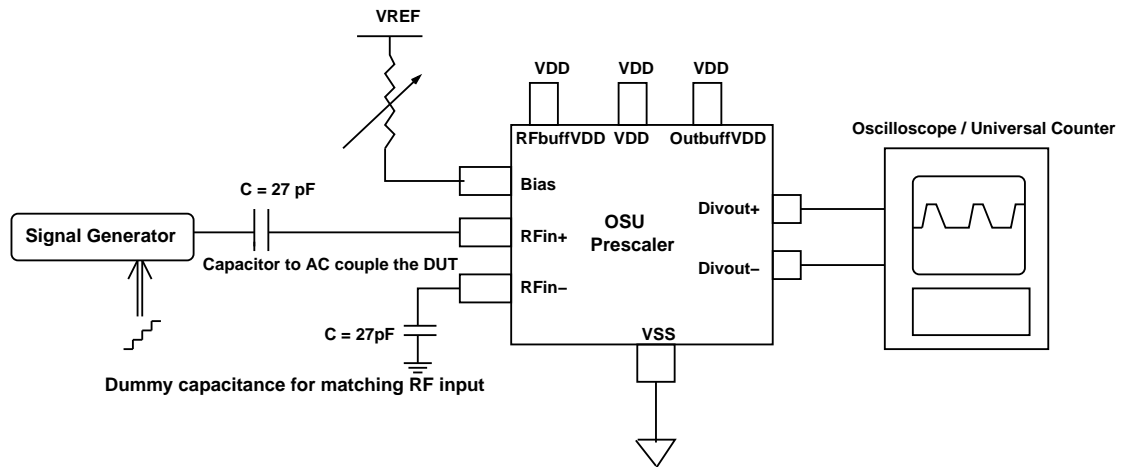

FIGURE 5.6: Test setup.

5.3.   Measurement Results

The measurement results of the dual-modulus prescaler operation have been tabulated in this section. The chip testing yielded the following measurement observations and inferences:

- The prescaler operation frequency varies with input power levels. The characteristic of the divider speeds when input levels are varied from -25 dBm to 0 dBm (corresponding to signal amplitudes of 18 mV – 320 mV) has been obtained. As shown in Figure 5.7, low power levels are likely to have lesser operation frequency because of the low noise margins. Very low power, high-frequency signals may be amplified insufficiently by the RF buffer inducing metastability in CML latches of the prescaler. At higher power levels ($\geq$ -16 dBm), corresponding to higher amplitudes, the second stage of the bipolar buffer may be driven into saturation. As bipolar transistors take longer time to get out of saturation, this could be the reason for lower clock speeds. Note that this implies the prescaler could still be capable of functioning and the dual-modulus operation is restricted by the buffer. Optimum prescaler speeds were obtained at signal amplitudes of $\approx$ 50 mV (-16 dBm).

- The maximum operating frequencies obtained from 21 chip samples and the optimum power-levels and biasing currents at which this was obtained are tabulated in Table 5.1.

- The output waveform is a divide-by-8/divide-by-9 alternating waveform which has been observed operational until frequency of 2.1 GHz with about 2 mA consumed in the CMOS prescaler. Figure 5.8 shows the output waveform as observed on the oscilloscope, with the specified biasing and signal conditions.

- When the bias current is set such that the current mirrored through the buffers

TABLE 5.1: Chip performance over 21 samples.

Sample Number	Max. frequency (GHz)	Signal amplitude (dBm)	Prescaler I (mA)
1	2.06	-16	2.10
2	1.85	-16	2.15
3	1.92	-15	2.10
4	2.00	-16	2.00
5	2.00	-23	2.30
6	2.06	-15	2.15
7	1.96	-23	1.96
8	1.81	-16	1.94
9	2.00	-15	2.45
10	2.00	-16	2.27
11	1.97	-16	2.35
12	2.03	-19	2.31
13	1.8	-16	1.96
14	2.10	-15	2.33
15	2.11	-20	2.38
16	1.85	-20	1.95
17	1.86	-20	1.78
18	1.8	-21	1.75
19	1.94	-23	2.06
20	1.9	-19	2.10
21	2.06	-17	2.46

Maximum prescaler operating frequency mean = 1.96 GHz

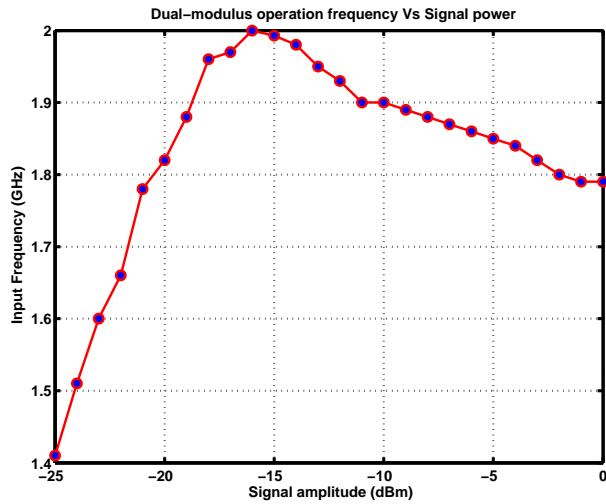Standard deviation $\sigma = 0.10$ GHz

FIGURE 5.7: Operating frequency variation with input signal levels.
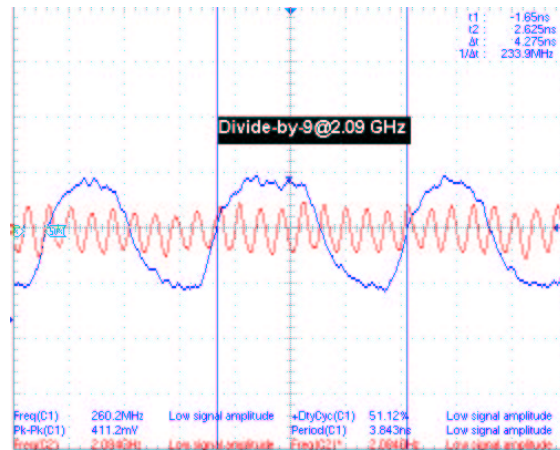
FIGURE 5.8: Measured operation at 2.1 GHz, 2.1 mA prescaler current, -16dBm input signal.

and the prescaler is according to the design, the maximum obtainable dual-modulus operation frequency is 1.85 GHz (Figure 5.9).

- Since the operation seems limited by the RF buffer stage transistors saturating, an experiment was done with higher supply voltage levels for the buffer alone. The frequency of operation was found to increase by only 10 MHz when supply voltage was increased from 2.5 V to 3 V.
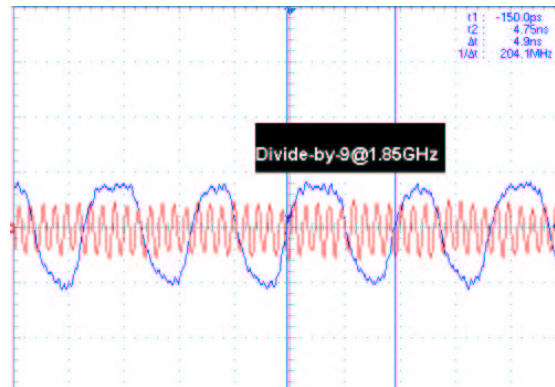
FIGURE 5.9: Measured operation at 1.85 GHz, 1.3 mA prescaler current, -16dBm input signal.

5.4.   Conclusions

An analysis of high-performance dual-modulus prescalers has been presented in this thesis. A design methodology was developed to implement high-speed dividers consuming lowest possible power. The system level issues of the pulse-swallow topology were investigated and modifications incorporated to optimize the propagation delays. The circuit was designed keeping in mind robustness to extremes of process and temperatures. The performance has been validated with measurements on a test chip fabricated with National BiCMOS8 0.25μm, 5-metal technology.

Future work on this topic could involve investigating other dual-modulus architectures like those in [28] and obtain perform comparisons. The preliminary results of this design could be used to design and fabricate the entire frequency synthesizer as that could open up several research avenues based on the influence of the various blocks in the mixed-signal chip.

BIBLIOGRAPHY

1. B. Razavi, *RF Microelectronics*, New Jersey: Prentice Hall, 1998.

2. L.Dai, *Design of High-Performance CMOS Voltage Controlled Oscillators*, Ph.D. thesis, University of Minnesota, Department of Electrical and Computer Engineering, 1996.

3. T. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge,UK: Cambridge University Press, 1998.

4. L.Lin, *Design Techniques for High-Performance Frequency Synthesizers for Multi-Standard Wireless Communication Applications*, Ph.D. thesis, University of California, Department of Electrical Engineering and Computer Science, 2000.

5. A. Hajimiri and T. Lee, "A general theory of phase noise in electrical oscillators," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 2, pp. 179–194, February 1998.

6. U. Moon, K. Mayaram, and J. Stonick, "Spectral analysis of time-domain phase jitter measurements," *IEEE Transactions on Circuits and Systems–II: Analog and Digital Signal Processing*, vol. 49, no. 6, pp. 321–327, May 2002.

7. T.A.D. Riley, M.A. Copeland, and T.A. Kwasniewsky, "Sigma-Delta modulation in fractional-N frequency synthesis," *IEEE Journal of Solid-State Circuits*, vol. 28, no. 1, pp. 553–559, May 1993.

8. K. Shu, E. Sanchez-Sinencio, F. Maloberti, and U. Eduri, "A comparitive study of digital sigma-delta modulators for fractional-N synthesis," in *Proc. of IEEE International Conf. Electronics,Circuits and Systems*, September 2001, pp. 1391–1394.

9. F. Gardner, "Charge pump phase-lock loops," *IEEE Transactions on Communications*, vol. COM-28, pp. 1849–1858, November 1980.

10. J. Hein and J. Scott, "Z-domain model for discrete-time PLLs," *IEEE Transactions on Circuits and Systems–II: Analog and Digital Signal Processing*, vol. 35, no. 6, pp. 1393–1400, November 1988.

11. H. Knapp, J. Bock, M. Wurzer, G. Ritzberger, K. Aufinger, and L. Treitinger, "2 GHz/2 mW and 12 GHz/ 30 mW dual-modulus prescalers in silicon bipolar technology," in *Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting*, September 2000, pp. 164–167.

12. L. Tournier, M. Sie, and J. Graffeuil, "A 14.5 GHz, 0.35-micron frequency divider for dual-modulus prescaler," in *IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, June 2002, pp. 227–230.

13. B.-U Klepser, "SiGe Bipolar 5.5 GHz dual-modulus prescaler," *IEE Electronics Letters*, vol. 35, no. 20, pp. 1728–1730, September 1999.

14. T.S. Aytur and B. Razavi, "A 2-GHz, 6-mW BiCMOS frequency synthesizer," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1457–1462, December 1995.

15. X.Li, *Evaluation of Radio Frequency CMOS Integrated Circuit Technology for Wireless LAN Applications*, Ph.D. thesis, University of Florida, Department of Electrical and Computer Engineering, 2003.

16. E. Abou-Allam, T. Manku, T. Ting, and M. Obrecht, "Impact of technology scaling on CMOS RF devices and circuits," in *IEEE Custom Integrated Circuits Conference*, May 2000, vol. 1, pp. 361–364.

17. B. Gilbert, "Why bipolar? :- http://www.chipcenter.com/analog/c009.htm," *Analog Avenue Columns,ChipCenter Electronics Group*, April 1998.

18. B. Razavi, *Design of Analog CMOS Integrated Circuits*, New York: McGraw-Hill, 2001.

19. A. Hastings, *The Art of Analog Layout*, New Jersey: Prentice Hall, 2001.

20. D.J. Allstot, S.-H Chee, S. Kiaei, and M. Shrivastawa, "Folded source-coupled logic vs. CMOS static logic for low-noise mixed-signal ics," *IEEE Transactions on Circuits and Systems–I: Fundamental Theory and Applications*, vol. 40, no. 9, pp. 553–563, September 1993.

21. Jason Musicer, *An Analysis of MOS Current Mode Logic for Low Power and High Performance Digital Logic*, Ph.D. thesis, University of California, Department of Electrical Engineering and Computer Science, 2000.

22. M. Mizuno, M. Yamashina, K. Furuta, H. Igura, H. Abiko, K. Okabe, A. Ono, and H. Yamada, "A GHz MOS adaptive pipeline technique using MOS current-mode logic," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 6, pp. 784–791, June 1996.

23. M. Alioto, G. Palumbo, and S. Pennisi, "Delay estimation of SCL gates with output buffer," in *Proc. of IEEE International Conf. Electronics,Circuits and Systems*, September 2001, pp. 719–722.

24. J. Craninckx and M. Steyaert, "A 1.75-GHz, 3-V dual-modulus divide-by-128/129 prescaler in 0.7-micron CMOS," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 7, pp. 890–897, July 1996.

25. J. Rabaey, *Digital Integrated Circuits : A Design Perspective*, New Jersey: Prentice Hall, 1996.

26. D. A. Johns and K. Martin, *Analog Integrated Circuit Design*, New York: John Wiley and Sons, 1997.

27. C-Y. Yang and Member S-I. Liu, "Fast-switching frequency synthesizer with a discriminator-aided phase detector," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 10, pp. 1445–1452, October 2000.

28. K. Shu, E. Sanchez-Sinencio, J. Silva-Martinez, and S.H.K. Embabi, "A 2.4-GHz monolithic fractional-N frequency synthesizer with robust phase-switching prescaler and loop capacitance multiplier," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 6, pp. 866–874, June 2003.