Charge-Steering Latches

by

Jacob Pike, B.Eng.

A thesis submitted to the
Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Master of Applied Science in Electrical Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
Department of Electronics
Carleton University
Ottawa, Ontario
July, 2018

©Copyright
Jacob Pike, 2018

Abstract

Modern wireline transceivers are explored to understand the limitations in their design and how their performance - in terms of speed and power consumption - can be pushed. This exploration leads to the revelation that charge-steering based transceivers have a high potential. As such, an in-depth transient analysis of the basic charge-steering latch is presented, providing new insight into the operation of charge-steering circuits. This insight is utilized to develop design equations for the basic charge-steering latch. The design equations lead to improvement in charge-steering latch design (lower power consumption, higher swing) when compared to prior art. While prior art focused on charge-steering systems (ie. data converters and SerDes), the focus on individual latch improvement could also improve these systems. Several new charge-steering latch topologies are also proposed. The transient analysis is extended to one of these latches, the charge-steering latch with a clocked cross-coupled pair. New design equations are also presented for this latch. Post-layout simulations of these latches are done in 28nm FD-SOI CMOS and further simulations are performed in 65nm CMOS. The new design equations and latch topologies enable the use of charge-steering latches at 28Gb/s and beyond, at new levels of performance. Power savings of as much as 40% are demonstrated by the new topologies, all other specifications being equal. As data transmission speeds are pushed to as much as 1 terabit per second, single lane transceivers are required which operate at 56Gb/s and beyond. The new charge-steering circuits and design equations can enable a 56Gb/s transceiver when used in a half-rate architecture.

Acknowledgments

First, I would like to acknowledge Dr. Calvin Plett for his supervision of my thesis. His willingness to supervise me almost a year after I started my Masters is very much so appreciated and his expertise of integrated circuits has been extremely valuable in the completion of my thesis. I also appreciate his facilitating my relationship with Ciena Corporation where I had the opportunity to learn from many other experts in integrated circuit design.

My manager at Ciena, Dr. Naim Ben-Hamida, has been very graceful in lending me his time, as have others on the Analog Integrated Circuit team, including Dr. Jerry Lam and Dr. Sadok Aouini. But a special thanks must go to Dr. Mahdi Parvizi. He deserves as much credit as anyone for this thesis. Not only did he point me in the initial direction of researching charge-steering circuits but he has spent countless hours looking at these circuits with me and giving me new ideas. I could not have completed this thesis without his help.

Many others at Ciena also deserve acknowledgment; Mouadh Abidi and Tigran Zohrabayan for their advice on layout and the other graduate students here, Matthew Mikkelson and Alex Jiang, for welcoming me into Ciena. I look forward to continuing my research with all of the above listed.

I must also acknowledge David Berton, who was instrumental in helping me through graduate courses and getting me started in learning Cadence. Lastly, I would be remiss not to acknowledge some of the other great people I met at Carleton along the way: Carmen Lefebvre, Mat with one T, Pablo Cortens and Joel Riordan. Saturdays are for the boys. Sundays are for Carmen.

Table of Contents

Abstract	ii
Acknowledgments	iii
Table of Contents	iv
List of Tables	vi
List of Figures	vii
Nomenclature	x
1 Introduction	1
1.1 Motivation	1
1.2 Contributions	4
1.3 Organization	4
2 Background	5
2.1 Conventional Wireline Transceivers	5
2.2 A Charge-Steering Based Receiver	9
2.3 Basic Charge-Steering Latch Versus Current-Steering Latch	11
2.4 StrongARM Latch	16
2.5 Other Conventional Charge-Steering Latches	17
2.6 A Comparison of Current-Steering and Charge-Steering Decision Feed- back Equalizers	20
3 New Charge-Steering Latch Analysis	23
3.1 Charge-Steering Latch In-Depth Analysis	23
3.1.1 Modes of Operation	24

3.1.2	Performance of the Basic Charge-Steering Latch	32
3.2	Design of the Basic Charge-Steering Latch	32
3.3	Charge-Steering with pMOS Cross-Coupled Pair	40
3.4	Charge-Steering with Clocked Cross-Coupled Pair	42
3.5	High-Swing Latches with Clocked Cross-Coupled Pair	45
3.6	High-Swing Latches with No Secondary Tail Capacitor	46
3.7	Charge-Steering Latch with Output Switch	47
4	Simulation of Charge-Steering Latches	50
4.1	Simulation of Basic Charge-Steering Latch	55
4.1.1	Basic Charge-Steering Latch 28nm Simulations	56
4.1.2	Basic Charge-Steering Latch 65nm Simulations	59
4.2	Comparison of Low-Power Charge-Steering Latches	61
4.2.1	Effect of Clock Buffers on Charge-Steering Latches	66
4.2.2	Charge-Steering Latch with Clocked Cross-Coupled Pair 28nm Simulations	68
4.3	Simulation of High-Swing Latches	69
4.4	Simulation of Charge-Steering Latch with Output Switch	73
5	Conclusion	78
5.1	Future Works	79
	List of References	81

List of Tables

4.1	Basic latch component sizes for sweeping simulations in 28nm	56
4.2	Basic latch component sizes from design equations in 28nm	57
4.3	Basic latch component sizes for sweeping simulations in 65nm	59
4.4	Basic latch component sizes from design equations in 65nm	59
4.5	Basic latch component sizes comparison in 65nm	61
4.6	Charge-steering latch with clocked cross-coupled pair component sizes for sweeping simulations in 28nm	68

List of Figures

1.1	Wireline transceiver block level overview.	2
1.2	Wireline transceiver applications and baseline specifications [3]. . . .	2
1.3	Single-lane wireline transceiver data rate trends vs. process node and year [8].	3
2.1	Typical channel loss for various wireline transceiver applications [3]. .	6
2.2	Continuous time linear equalizer with inductive peaking.	7
2.3	Traditional feed forward equalizer architecture.	7
2.4	Traditional decision feedback equalizer architecture.	8
2.5	Bang bang phase detector.	9
2.6	Full receiver architecture from [16].	10
2.7	Non-return to zero charge-steering latch.	11
2.8	Half-rate bang bang phase detector.	12
2.9	Basic current-steering latch.	12
2.10	Basic charge-steering latch.	13
2.11	Charge-steering latch versus current-steering latch transient operation.	15
2.12	Original StrongARM latch.	17
2.13	Charge-steering latch with pMOS cross-coupled pair.	18
2.14	High-swing charge-steering latch 1.	18
2.15	High-swing charge-steering latch 2. Has the addition of cross-coupling the cascode devices.	19
2.16	Single tap current-mode logic decision feedback equalizer.	21
2.17	Half-rate decision feedback equalizer path from [16].	21
3.1	Basic charge-steering latch transient operation.	25
3.2	Charge-steering latch model during reset.	26
3.3	Charge-steering latch during Mode 2. (a) Full model. (b) Simplified model.	27
3.4	Charge-steering latch model during evaluation.	29

3.5	Charge-steering latch model transient operation during evaluation. . .	30
3.6	Latch predicted swing for a given tail capacitor to output capacitor ratio. . .	34
3.7	Charge-steering latch with clocked pMOS cross-coupled pair.	42
3.8	Charge-steering latch comparison.	43
3.9	High-swing charge-steering latch with clocked pMOS cross-coupled pair. . .	45
3.10	High-swing charge-steering latch with clocked pMOS cross-coupled pair and cross-coupled cascode devices.	46
3.11	High-swing charge-steering latch with no secondary tail capacitor. . .	47
3.12	High-swing charge-steering latch with no secondary tail capacitor and cross-coupled cascode devices.	48
3.13	Basic charge-steering latch with output switch.	49
4.1	Basic charge-steering latch transient operation in worst case scenario, demonstrating one of the input differential pair transistors never turns off during evaluation.	52
4.2	Basic charge-steering latch equivalent model during Mode 2. a) Full model. b) Simplified model.	53
4.3	Voltage on node P during Mode 2 of basic charge-steering latch model. (a) 5 Gb/s. (b) 10 Gb/s. (c) 14 Gb/s. (d) 20 Gb/s. (e) 28 Gb/s. . .	54
4.4	Clock skew effect on basic charge-steering latch swing and power con- sumption.	55
4.5	Layout for basic charge-steering latch in 28nm CMOS.	57
4.6	Figure of merit comparison for basic latch in 28nm.	57
4.7	Figure of merit for basic latch in 28nm for various input transistor widths.	58
4.8	Figure of merit comparison for basic latch in 65nm.	60
4.9	Performance comparison with prior art in 65nm.	60
4.10	Basic charge-steering latch transient operation.	62
4.11	Critical amplitude versus power consumption for low-power latches. (a) 5Gb/s data rate. (b) 10Gb/s. (c) 14Gb/s. (d) 20Gb/s. (e) 28Gb/s. . .	64
4.12	Corner simulations for charge-steering latch with a clocked cross- coupled pair. (a) 5Gb/s data rate. (b) 10Gb/s. (c) 14Gb/s. (d) 20Gb/s. (e) 28Gb/s.	65
4.13	Power consumption at peak amplitude of low-power latches with com- parison to predicted power savings.	66

4.14	Power consumption at peak amplitude of low-power latches including power added by clock buffers.	67
4.15	Layout for charge-steering latch with clocked cross-coupled pair in 28nm CMOS.	69
4.16	Figure of merit comparison for charge-steering latch with clocked cross-coupled pair in 28nm.	69
4.17	Critical amplitude versus power consumption for high-swing latches. (a) 5Gb/s data rate. (b) 10Gb/s. (c) 14Gb/s. (d) 20Gb/s. (e) 28Gb/s.	71
4.18	Power consumption at peak amplitude of high-swing latches.	72
4.19	Basic charge-steering latch and charge-steering latch with output switch transient simulations at 10Gb/s.	75
4.20	Power consumption of basic charge-steering latch versus charge-steering latch with output switch.	76
4.21	Total width of all precharge switches in the basic charge-steering latch versus charge-steering latch with output switch.	77

Nomenclature

Gb/s	Gigabit per Second
Tb/s	Terabit per Second
SerDes	Serializer-Deserializer
SR	Short Reach
MR	Medium Reach
LR	Long Reach
PCB	Printed Circuit Board
CEI	Comitato Elettrotecnico Italiano or Italian Electrotechnical Committee
IO	Input Output
mW/Gb/s	Milliwatt per Gigabit per Second
pJ/bit	Picojoule per Bit
BER	Bit Error Rate
CS	Charge-Steering
GHz	Gigahertz
NRZ	Non-Return to Zero
PAM4	Pulse Amplitude Modulation with 4 Amplitude Levels
ISI	Inter-Symbol Interference
FFE	Feed Forward Equalizer
DFE	Decision Feedback Equalizer
CTLE	Continuous Time Linear Equalizer
UI	Unit Interval
CML	Current-Mode Logic
CDR	Clock and Data Recovery
VCO	Voltage Controlled Oscillator
PD	Phase Detector
BBPD	Bang Bang Phase Detector
MUX	Multiplexer

DMUX	Demultiplexer
DTLE	Discrete Time Linear Equalizer
CC	Cross-Coupled
RZ	Return-to-Zero
FoM	Figure of Merit
FS	Fast nMOS-Slow pMOS
SF	Slow nMOS-Fast pMOS
FF	Fast nMOS-Fast pMOS
SS	Slow nMOS-Slow pMOS
TT	Typical nMOS-Typical pMOS
PLL	Phase-Locked Loop
ILO	Injection Locked Oscillator
UTBB	Ultra Thin Body and Buried Oxide
FD-SOI	Fully Depleted Silicon on Insulator
CMOS	Complementary Metal Oxide Semiconductor

Chapter 1

Introduction

Chapter 1.1 will begin by discussing data transmission in modern technology and briefly look at the future of it, including what speeds, channel losses and power consumption are expected of data transmission circuitry in various applications. Following this, in Ch. 1.2, contributions made by this thesis to modern transceiver design will be reviewed. Finally, Ch. 1.3 will present the organization of this thesis.

1.1 Motivation

As societies reliance on communicating via technology continues to increase, data transmission speeds must also increase. Each year consumer demands require that our electronics - phones, computers, servers, et cetera - get faster, requiring the processors, memory units and other chips within, communicate faster. Currently there are wireline data transmission standards up to 400 gigabits per second (Gb/s) [1]. Shortly after 2020 there are expected to be standards set for 1 terabit per second (Tb/s) and beyond [2]. In order to transmit and receive data at these speeds, hardware that can operate at these speeds will be required.

Transmitters and receivers - collectively known as a transceiver or Serializer-Deserialzer (SerDes) - is the key component of the hardware which enables data transmission. A typical wireline transceiver block diagram can be seen in Fig. 1.1. The basic function of the transmitter is to serialize many parallel streams to then be sent across a channel. The receiver then recovers the incoming serial data. The complexity of the transmitter and receiver varies from application to application and is dictated by, among other things, the channel. Transceivers are used in many applications but are split into three main categories; Short Reach (SR), Medium Reach

Figure 1.1: Wireline transceiver block level overview.

Figure 1.2: Wireline transceiver applications and baseline specifications [3].

(MR) and Long Reach (LR). SR applications include data transmission from chip-to-module across very short Printed Circuit Board (PCB) traces, where there is very little loss. MR applications include data transmission from chip-to-chip across PCB traces less than 50cm in length or through mezzanines. LR applications are from chip-to-chip, where both chips are on entirely different PCBs, through a copper cable or a backplane. Figure 1.2 illustrates these three applications and shows the losses which can be expected across the channel in each application according to the Italian Electrotechnical Committee (CEI), which sets standards for single-lane transceivers.

In order to enable data transmission at higher speeds, several single-lane

Figure 1.3: Single-lane wireline transceiver data rate trends vs. process node and year [8].

transceivers are used in parallel and together they make up a multi-lane SerDes. However, as the feature sizes of integrated circuits continue to decrease, the number of Inputs and Outputs (IOs) to a chip must also decrease. This demands that less single-lane transceivers may be used in parallel in a multi-lane SerDes. Figure 1.3 shows the speed of state of the art single-lane wireline transceivers over the last decade. It can be seen that research is trending towards 56Gb/s and 112Gb/s transceiver design in order to enable data transmission at 400Gb/s and 1Tb/s. While several recent designs have demonstrated the ability to operate in this speed range [4]–[7], this is only one of the criteria these transceivers must meet.

The performance of transmitters and receivers is often quantified by their power efficiency given in milliwatts per gigabit per second (mW/Gb/s) or picojoules per bit (pJ/Bit). The power efficiency of transceivers must decrease as their speed increases in order to keep the total power consumption relatively constant. Recent designs have targeted a power efficiency of 1mW/Gb/s [9]–[12]. While there are other performance metrics - channel loss at Nyquist, die size and Bit Error Rate (BER) for example - power efficiency and speed are perhaps the most important to transceiver design.

While no reported transceiver has achieved a speed of 56Gb/s and efficiency of

1mW/Gb/s, there are many promising designs. Among the most promising of these designs are receivers based on charge-steering (CS) techniques [13]–[18]. Analyzing and improving CS circuits is the focus of this thesis. By understanding and improving CS circuits, the hope is to enable a 56Gb/s transceiver at power efficiencies improving upon prior art.

1.2 Contributions

This thesis presents a new analysis and design methodology for optimizing CS latches. The new understanding leads to the development of several new CS topologies, which are also presented. The new topologies and design equations demonstrate power savings, increased swing and/or the ability for CS latches to operate at higher speeds. [16] used CS latches in a 40Gb/s, 14mW transceiver, demonstrating the ability of CS latches to operate at 20Gb/s. This work demonstrates that CS latches can be used at 28Gb/s and beyond enabling a 56Gb/s wireline transceiver when used in a half-rate architecture. Conversely, the work demonstrates the ability to have a 40Gb/s transceiver with a power below 14mW when the new CS latch topologies and design methodology are used.

1.3 Organization

This thesis is organized as follows. Chapter 2 briefly discusses some conventional wireline receiver topologies and some more details on the receiver presented in [16] to demonstrate how crucial CS latches are to its design. Chapter 2 continues on by analyzing and comparing conventional CS circuits to more commonly used latches. Chapter 3 analyzes the transient operation of the basic CS latch in depth and presents new design equations for this latch. This analysis leads to some new CS topologies, which are also presented in Ch. 3. Chapter 4 presents the layout of several charge-steering latches at several data rates. Post-layout simulations results are presented, which demonstrate the efficacy of the new design equations and topologies. Chapter 5 concludes the work and presents some ideas for future work.

Chapter 2

Background

Chapter 2.1 will begin by exploring the challenges facing wireline transceiver design as speeds are pushed, and power efficiency and channel loss become defining specifications in design. The various methods for equalizing data and recovering clocks in serial link design are explored. Modern implementations of equalization and clock circuitry is explored, including looking at Current-Mode Logic (CML). While, Ch. 2.2 looks at CS based implementations of equalization and clock recovery circuitry. A comparison of CS and CML is then performed in Ch. 2.3. Chapter 2.3 also provides more specifics on the operation of the basic CS and CML latch. Chapter 2.5 does the same for more complex implementations of CS latches. With all conventional CS latches understood, Ch. 2.6 returns to looking at equalization systems, in particular the Decision Feedback Equalizer (DFE), which helps the reader to fully understand the importance of latches in wireline transceivers.

2.1 Conventional Wireline Transceivers

As data transmission speeds continue to increase, the need for more complex transmitters and receivers follows. This is primarily due to the insertion loss across a channel. The typical channel profile for the three main channel types (SR, MR and LR) can be seen in Fig. 2.1. As data rates continue to increase, the insertion loss begins to worsen. If data is sent at 56Gb/s, then the data could have a frequency profile up to 28 Gigahertz (GHz) using a Non-Return to Zero (NRZ) encoding or 14GHz using a Pulse Amplitude Modulation with 4 Amplitude Levels (PAM4) encoding. It can be seen that, even for SR applications, the channel loss at these frequencies is -23dB and -10dB respectively. This causes amplitude attenuation of as much as 1/10th

Figure 2.1: Typical channel loss for various wireline transceiver applications [3].

and 1/3rd as well as Intersymbol Interference (ISI), where bits disperse into the bits preceding and following it.

The issue of amplitude attenuation and ISI are addressed using equalizers. There are three main types of equalizers; a Feed Forward Equalizer (FFE), DFEs and a Continuous Time Linear Equalizer (CTLE). A typical CTLE can be seen in Fig. 2.2. CTLEs are used on the receive side and are analog amplifiers, which help to increase amplitude after data has passed through a channel, but they cannot correct for ISI. Conversely, an FFE is most commonly used in the transmit side and can correct for precursor ISI. A typical FFE block diagram can be seen in Fig. 2.3. Incoming data is delayed N times through latches and after each stage of delay the sampled data is used to adjust a tap coefficient. All the tap coefficients are then summed with the last stage of delayed data. The taps will add or take away from the current bit to prevent it from spreading into other bits as it passes through the channel. The number of delay stages or taps that are required is dictated by the pulse response of the channel that is being used. A channel with a bad pulse response will see the current bit spread across many Unit Intervals (UIs), into several bits before and after it. More taps are required to correct for this.

DFEs are used on the receive side and correct for post-cursor ISI. They are the

Figure 2.2: Continuous time linear equalizer with inductive peaking.

Figure 2.3: Traditional feed forward equalizer architecture.

Figure 2.4: Traditional decision feedback equalizer architecture.

most crucial equalizer in SerDes design. A typical DFE block diagram can be seen in Fig. 2.4. DFEs work similarly to FFEs, though in reverse. Bits preceding the current one are sampled and used to adjust the tap coefficients as to emphasize or deemphasize the current bit and correct for post-cursor ISI. Again, the number of taps required depends on the channel pulse response. Many traditional FFEs and DFEs use CML latches. CML latches and CML based equalizers will be explored in more in depth later in this Chapter, but regardless of how they are implemented it is clear that latches form the foundation of these equalizers.

The primary goal of a SerDes is to minimize the number of IOs to a chip and as such the number of wires into and out of a chip. In order to accomplish this goal a clock is not transmitted to the receive side of a channel. As such a Clock and Data Recovery (CDR) circuit is required to recover the clock which corresponds to the data transmission speed. A typical CDR consists of a loop filter and a Voltage Controlled Oscillator (VCO), but the crucial component is the Phase Detector (PD). The PD measures the phase difference between the incoming data and the recovered clock. The PD then drives the loop filter, which generates a reference voltage for the VCO, which in turn adjusts the recovered clock to align with the incoming data. In NRZ encoded SerDes a two times oversampling Bang-Bang (or Alexander) Phase Detector (BBPD) is often used as seen in Fig. 2.5. In a BBPD, the incoming data is sampled three times per clock period. If the first two samples differ from the last sample, the clock is late. If the first sample differs from the last two samples, the clock is early. If all samples are the same, there has been no transition in the data. The output of

Figure 2.5: Bang bang phase detector.

the two XOR gates will reflect this and will drive the VCO accordingly. It is clear that latches again make up the majority of the circuitry in a BBPD. Other typical PDs, for the example Hogge PDs and other linear PDs, are similarly latch based.

While there are other important components to discuss - for example the multiplexing (MUXing) in the transmitter to serialize the many parallel streams - the CDR and equalizer circuitry is most important in SerDes design. While many traditional SerDes are based on CML, recent designs have demonstrated the benefits of CS latches in SerDes applications.

2.2 A Charge-Steering Based Receiver

CS latches were first presented in [19], which dates to 2012, but updated receivers have been presented in recent years. The focus of [19] was on the CDR, however in [15] - and expanded upon in [16] and [20] - a full receiver is presented with equalization and a CDR circuit. This receiver achieves a BER of less than 10^{-12} while recovering data at 40Gb/s with a power dissipation of 14mW, for a power efficiency of 0.35mW/Gb/s. To the knowledge of the author of this thesis, no receiver has achieved a comparable power efficiency at such a high-speed. The full receiver architecture can be seen in Fig. 2.6.

First, the equalizer architecture of this receiver will be discussed. The first stage in the equalizer is a traditional CTLE used to gain some amplitude in the incoming data.

Figure 2.6: Full receiver architecture from [16].

Following this a Demultiplexer (DMUX) is used to split the data into an even and odd data stream. The DMUX is made up of two NRZ CS latches as seen in Fig. 2.7. Following the DMUX is the Discrete Time Linear Equalizer (DTLE). The DTLE works similar to an FFE in the transmitter and performs precursor de-emphasis. However, this particular DTLE performs the sampling using a switched-capacitor as opposed to a flip flop or latch. The heart of the equalization is the DFE. This receiver uses a half-rate, two-tap DFE adapted from the traditional DFE. The DFE is made up of a two summers, four MUXs and eight latches, all of which are based on CS latches, illustrating the importance of their design to the overall architecture. The DFE in this architecture will be explored more in depth later in this Chapter when CS circuits have been explored as well.

The other half of the receiver in Fig. 2.6 is the CDR. This receiver's CDR employs a linear BBPD as seen in Fig. 2.8. This PD was first presented in [21]. The first two latches and output of their XOR generate the phase error signal. To eliminate data dependency issues, which arise in Hogge PDs, the last four latches and their XOR generate a constant pulse width every time there is a data transition. The reference pulse is then taken away from the error pulse to generate an unambiguous error signal. Again, it can be seen that the PD is almost entirely made up of latches, which in this

Figure 2.7: Non-return to zero charge-steering latch.

case are all CS circuits and further illustrates the importance of CS latches to this receiver.

In order to enable the 40Gb/s receiver, the CS latches which make it up need to operate as high as 20Gb/s, given the CDR and DFE operate at half-rate. Thus in order to enable a 56Gb/s receiver, it is likely that 28Gb/s latches will be required. There are many different types of CS latches presented in [16] but first the basic CS latch and their benefit over CML latches must be understood.

2.3 Basic Charge-Steering Latch Versus Current-Steering Latch

As previously mentioned conventional high-speed transceivers utilize current-steering based circuits, also known CML. A basic CML latch can be seen in Fig. 2.9. In the tail of the latch there is a constant current source, which is steered to the load resistors via an input differential pair. When the input to one transistor is high, that transistor will be conducting strongly and the current will be steered towards the corresponding resistor, thus creating a voltage at the output. The other transistor will have little conduction and less current will be steered towards that branch, so a smaller voltage will appear at the drain of that transistor. This creates a differential

Figure 2.8: Half-rate bang bang phase detector.

Figure 2.9: Basic current-steering latch.

swing at the output. While CML latches can operate at high-speeds, they always consume power due to the constant current source in the tail. This makes it difficult for a CML based transceiver to achieve a good power efficiency.

The basic CS latch can be seen in Fig. 2.10. As described in [22], the basic latch consists of a differential input pair (M_1/M_2), four switches ($M_3/M_4/M_5/M_6$) and a tail capacitor (C_T). Furthermore there is a capacitor (C_D) at each of the two output nodes. These capacitors can be extrinsically defined or can simply be the input impedance of a following stage.

Previously the basic CS latch has been described as having three crucial nodes: the output nodes (X/Y) and the tail node (T). Furthermore, CS latches have been

Figure 2.10: Basic charge-steering latch.

described as having two modes of operation: reset and evaluation. In reset mode, the input clock is low such that M_3/M_4 and M_6 are ON. This allows C_D to charge to V_{DD} and C_T to discharge to ground.

When the input clock goes high, M_3/M_4 and M_6 are OFF, M_5 is ON and the latch enters into evaluation mode. One input transistor, for example M_2 , is OFF due to its input being low. At this point the other transistor, M_1 , is ON as the input to its gate is high. One output node, X , discharges across M_1 and M_5 to the tail capacitor. In ideal operation the other output node, Y , has no path to discharge and remains at V_{DD} . This creates a differential swing at the output which tracks the input until the circuit enters into reset mode again.

Following this analysis, the swing of the basic CS latch can be predicted as [19]

$$\Delta V_{CD} = V_{DD} - V_{DD} \frac{1}{1 + \frac{C_T}{C_D}} = V_{DD} \frac{C_T}{C_T + C_D}. \quad (2.1)$$

Furthermore, the current consumption of the basic latch can be predicted as [22]

$$I = \frac{C_D \Delta V_{CD}}{T_b} \quad (2.2)$$

where T_b is the clock period.

The following analysis is presented in [22] and is useful in understanding the basic

CS latch. In order to produce the same swing at the output of a CML latch, the current-resistor product of the CML latch must be equal to ΔV_{CD} such that

$$I_{SS} = \frac{\Delta V_{CD}}{R_D}. \quad (2.3)$$

It can also be said that the current-resistor product of the CML latch must produce a bandwidth which is 70% of the data rate such that the resistor is given by

$$R = \frac{T_B}{1.4\pi C_D}. \quad (2.4)$$

This tells us that the current consumption of a CML latch can be approximated by

$$I_{SS} = \frac{1.4\pi C_D \Delta V_{CD}}{T_B}. \quad (2.5)$$

Dividing (2.2) by (2.5), it can be seen that the CS latch should have a power reduction of about 4.4 times over a CML latch that produces the same output swing. Fig. 2.11 shows a transient comparison between the two latches. The simulation was performed with Cadence Spectre using STMicroelectronics's 28nm process, at a data rate of 10Gb/s, with a common-mode voltage of 850mV, an input with a 100mVp swing and sinusoidal clocks. It can be seen that both the CS latch and CML latch achieve the same differential output amplitude with the input pair in both set to the same size. The CS latch consumes about a quarter of the power ($169.2\mu A/750\mu A$) of the CML latch, as expected.

It can also be seen that the CS latch output has a Return-to-Zero (RZ) encoding due to the reset mode. In some applications this will need to be adjusted to NRZ, which can be accomplished using various track and hold circuits. However, in other applications where the setup time of the latch is crucial, the RZ nature of the CS latch is advantageous. For example in a SerDes the first tap in the DFE can at most have a delay through the summing node, latch and tap of 1UI in order to affect the desired bit. A CML latch has a setup time which is twice that of a CS latch since it must travel from its peak amplitude to the common mode level during evaluation,

Figure 2.11: Charge-steering latch versus current-steering latch transient operation.

whereas a CS latch begins evaluation having already settled to the common mode.

While CS latches have a much lower power consumption when compared to CML, they do require full-swing clocks to drive their switches. This issue is inherent in CS latch design and while it is certainly one of the main concerns, the full-swing clocks may be worth the significant power savings.

The maximum speed of operation of CS circuits may also be limited compared to CML. As the speed increases, the output capacitors of the CS latch does not have much time to discharge and at a high enough speed, the output would fall below unity gain. CML circuits do not face this issue and can operate at much higher speeds. However, this is not a concern for current single-lane SerDes speeds. CS circuits still see the fourfold power savings when used in SerDes operating at 56Gb/s and beyond, as the fundamental equations presented in this chapter demonstrated and as simulations will verify in Ch. 4.

2.4 StrongARM Latch

Another latch used in conventional SerDes design is the StrongARM latch (Fig. 2.11). The StrongARM latch shares many advantages with CS latches, including zero static power consumption. Furthermore, the StrongARM latch produces rail-to-rail outputs at lower speeds. However, the power consumption of the StrongARM latch is significantly higher. Similar to CS latches, the power is primarily consumed in charging and discharging capacitors each cycle, however additional devices ($M_4/M_5/M_6/M_7$) increase the output capacitance and rail-to-rail swings mean these capacitors must be charged further each period. Furthermore, the additional devices ($M_4/M_5/M_6/M_7$) limit the max speed of operation, such that the StrongARM latch may not operate in 56Gb/s or 112Gb/s SerDes. The StrongARM latch and how the addition of Cross-Coupled (CC) devices limits speed, will be explored more in Ch. 4, but for now it should be remarked that in lower speed applications (on the range of 10Gb/s) and where rail-to-rail outputs are required it can be an effective latch. But for speeds beyond 10Gb/s and in extremely low power applications, an alternative is needed.

While the basic CS latch forms the foundation of all CS circuits, there are other CS latches, which have been designed for further applications.

Figure 2.12: Original StrongARM latch [23].

2.5 Other Conventional Charge-Steering Latches

The second latch topology under consideration adds a pMOS cross-coupled pair (M_7/M_8) to the output nodes as seen in Fig. 2.13. As discussed previously, when the basic CS latch enters evaluation mode charge sharing occurs between both output nodes and C_T . Adding the pMOS CC pair can give one output node a path to V_{DD} during evaluation, which can increase the differential swing while reducing the combined single-ended swing. This results in power savings when charging the output capacitors during reset. However, the charging of one output node during evaluation cancels the power savings out.

While the CS latch with a CC pair provides a higher swing than the basic latch at a comparable power consumption, [16] also presents latches that can achieve a much higher swing at the cost of increased power consumption. The first of such latches can be seen in Fig. 2.14.

The addition of the pMOS CC pair restores one output level to V_{DD} and keeps the input transistors in saturation during evaluation mode. However, the output level can only be fully restored to V_{DD} given infinite time. Adding two CC cascode devices (M_{10}/M_{11}) isolates the drain of the input differential pair from the capacitance at the output. This delays the turn on of the unwanted input transistor. Charge

Figure 2.13: Charge-steering latch with pMOS cross-coupled pair.

Figure 2.14: High-swing charge-steering latch 1.

Figure 2.15: High-swing charge-steering latch 2. Has the addition of cross-coupling the cascode devices.

is therefore only transferred from one output node until the drain voltage of the unwanted transistor falls below $V_{DD} - V_{TH}$. Another CC nMOS pair (M_{12}/M_{13}) is added in the tail increasing the voltage at the drain of the differential pair, which further delays the turn on of the unwanted transistor. These extra devices increase the swing of the latch but at the cost of power consumption. The additional power consumption is difficult to predict but can be significant due to additional parasitic capacitances and increased discharge of the output capacitors. All of these capacitors will then be charged during reset and will consume more power in doing so.

Another high-swing latch can be seen in Fig. 2.15. M_{10} and M_{11} are now a CC pair, which helps to further delay the turn on of the unwanted input transistor, say M_2 . As the drain of the desired input transistor, M_1 , drops, M_{10} begin conducting more strongly. The voltage at the drain of M_{10} then begins dropping, which now lowers the gate voltage on M_{11} . This should help keep M_2 OFF for longer and prevent the unwanted discharge of the second output capacitor. Cross-coupling these transistors increases the parasitics at the output node, consumes more power and decreases the max speed of operation, but it can also result in a higher output swing.

Each of the CS latches presented in this Chapter have their place in different applications. The basic CS latch can operate very fast but provides a limited output swing. Adding the pMOS CC pair helps to extend the output swing without

significantly decreasing the potential speed of operation or increasing the power consumption. The two high-swing latches both significantly improve the output swing but cannot operate as fast as the former latches and consume much more power.

A new, more in depth exploration of CS latches will be presented in Ch. 3, but first we will return to looking at the traditional CML-based DFE and the CS-based DFE.

2.6 A Comparison of Current-Steering and Charge-Steering Decision Feedback Equalizers

With CML and CS latches better understood, we can now explore equalizers more in depth to further understand how the latches work as part of a system. Traditional FFEs and DFEs have been based on CML. A one-tap DFE based on CML can be seen in Fig. 2.16, with each component of the DFE identified. A bit is initially sampled by the summer. As explained previously, a CML-based summer will see the input into one gate be high at a time and the tail current will be steered towards this path thus creating a differential swing at the output. The summer output is then sent into another CML-based latch which again samples the data. This second sample is used to drive another differential pair, identified as the tap in Fig. 2.16. If a bit is high during the second sampling, current will be steered from the tap's tail current to the corresponding branch in the summer. This will help create a larger voltage at one output of the summer, in turn reducing its differential swing. This helps ensure the first sampled bit does not spread into the current bit, reducing post-cursor ISI. Conversely when a bit is low, both transistors in the tap will steer equal current to the current bit, having a negligible effect.

A single half-rate CS DFE path from [16] can be seen in Fig 2.17 and the similarity to the CML-based DFEs is evident. While the CS DFE may look more complicated, this is due to the addition of a second tap and the interleaved half-rate architecture that is employed. The CS DFE functions very similarly to the CML DFE. A bit is sampled by the summer and then two high-swing latches, identified as L_1 and L_2 , perform a second sampling. The results of the second sampling is again used to drive a differential pair in the tap, which in turns adds more capacitance to the

Figure 2.16: Single tap current-mode logic decision feedback equalizer.

output nodes of the summer as required. Every component of this DFE is based on CS latches. While the output of the CS DFE is RZ, the power savings of 4.4 times per latch over CML makes a CS-based DFE attractive.

Figure 2.17: Half-rate decision feedback equalizer path from [16].

Chapter 2.1 and 2.2 looked at the common methods - including CTLEs, DFEs, FFEs and BBPDs - for equalizing data and recovering clocks. This exploration revealed the importance of CML and CS latches to modern transceiver design. These two methods (current-steering and charge-steering) were explored more in depth in

Ch. 2.3. Chapter 2.5 looked at more complex CS latches. With CS latches understood, Ch. 2.6 provided specifics on the operation of DFEs. This understanding can easily be extended to FFEs and other common SerDes circuitry. While Ch. 2 provided some detail on CS latch operation, a robust design methodology for CS latches is required in order to optimally design them and CS based SerDes. Chapter 3 will look more in depth into the transient operation of these CS latches in order to develop such a methodology.

Chapter 3

New Charge-Steering Latch Analysis

Chapter 3 begins by looking at the transient operation of the basic CS latch, introducing several new considerations and modes of operation not seen in conventional analysis. This analysis is used to develop new design equations for the basic CS latch. Chapter 3.3 continues this transient analysis and design for the CS latch with a pMOS CC pair. This analysis leads to several new topologies presented in Ch. 3.4, 3.5, 3.6 and 3.7.

3.1 Charge-Steering Latch In-Depth Analysis

To improve our understanding of the basic CS latch it is important to consider a fourth node, the source of the differential pair (node P), in the transient analysis. There are parasitic capacitances from M_1/M_2 and M_5 at node P . The total parasitic capacitor, C_P , is charged towards V_{DD} during reset and discharged onto the tail during evaluation, similarly to C_D . This node will not completely charge to V_{DD} during reset because its path to the supply also sees the resistance of M_1/M_2 , in addition to being charged through M_3/M_4 . However, the swing at this node is significant enough to receive consideration when discussing the performance of a CS latch.

Furthermore, it cannot be stated that during evaluation only one input transistor is ON at a time. While with a small common-mode voltage and large input swing this may be true, the intended use of CS latches is in low power applications where input swing will not be sufficiently large. Furthermore, CS latches are often cascaded where the common-mode voltage is only in the designer's control for the first stage and is large in following stages. In actuality both transistors, M_1 and M_2 , will have some conduction. This results in both outputs discharging onto the tail at the same

time, though one much faster than the other.

Finally, it cannot be stated that there are only two modes of operation. In reality there will be four modes of operation: reset, evaluation and two in-between periods when the clocks overlap and all switches are ON and conducting. These modes will now be discussed.

3.1.1 Modes of Operation

The basic transient operation of a CS latch can be seen in Fig. 3.1. As before, the simulation was performed with Cadence Spectre using STMicroelectronics's 28nm process, at a data rate of 10Gb/s, with a common-mode voltage of 850mV, an input with a 100mVp swing and sinusoidal clocks. The four modes of operation are identified and to fully understand how CS latches operate, we must understand each of these modes, beginning with reset mode.

Mode 1: Reset

During the reset period, the four aforementioned nodes are charged/discharged to their final voltages. The latch configuration during this period can be seen in Fig. 3.2. It can be seen that C_{D1} , C_{D2} and C_P are charged by the supply voltage through some resistance equivalent to the drain-source resistance of M_1 - M_4 . It is known that the voltage on a charging capacitor in this configuration is given by

$$V(t) = V_{DD}[1 - e^{-\frac{t}{RC}}] \quad (3.1)$$

where t is the time for which the capacitor has been charging, R is the equivalent resistance between V_{DD} and the capacitor, C . If a capacitor has been charging for $t = 5RC$ then it will reach over 99% of the supply. If designed properly, M_3 and M_4 are conducting during reset and their effective resistance is small, allowing C_{D1} and C_{D2} to reach V_{DD} during reset.

Given infinite time, the final voltage on C_P would also be V_{DD} but the resistance between C_P and V_{DD} is large enough such that this voltage will never be reached. This is because C_P also sees the resistance of M_1 and M_2 , which will be conducting with weak to moderate inversion and will have a large equivalent resistance. Instead

Figure 3.1: Basic charge-steering latch transient operation.

Figure 3.2: Charge-steering latch model during reset.

the final voltage on C_P , the source of $M_{1,2}$, will approximately follow the gate voltage minus the process threshold voltage. This voltage has a slight dependence on the data rate as the equivalent resistance across M_1 - M_4 changes with frequency. However, the drain-source resistance of M_1/M_2 is much more dependent on their gate voltages, which is specified by V_{CM} and the input swing. Transistor sizes have a negligible effect on the voltage on this node as the parasitic capacitance at the node scales up as the resistance scales down, effectively creating the same time constant.

Like M_1 and M_2 , M_6 is conducting and therefore its drain-source resistance is small enough that the tail capacitor should fully discharge to 0V during reset. M_6 , if made small enough, can cause C_T to not fully discharge to ground but this should be avoided when designing a CS latch [19]. More details on effectively designing a CS latch will be provided later.

Mode 2: All Switches Closed

Immediately following reset, all switches will be closed for a short period of time. During this period all the capacitive nodes are connected. However, the resistance of M_1 and M_2 will be large enough so no charge-sharing will occur between C_{D1}/C_{D2} and the other capacitors. On the other hand, M_5 and M_6 will both be on during this period such that we have the latch configuration as seen in Fig. 3.3 (a). The exact

Figure 3.3: Charge-steering latch during Mode 2. (a) Full model. (b) Simplified model.

period of time during which all switches are closed depends on whether the clock is sinusoidal or square, what the rise and fall times of the clock are and the threshold voltage of the process being used. For a sinusoidal clock it can be predicted as

$$t_{mode2,sinusoidal} = 2 \left| \frac{\sin^{-1}\left(\frac{2V_{TH}}{V_{amp}} - 1\right)}{2\pi f} \right| \quad (3.2)$$

where V_{TH} is the process threshold voltage, V_{amp} is the clock amplitude and f is the clock frequency. For a square clock, presuming equal rise and fall times, it can be predicted as

$$t_{mode2,square} = \left(1 - \frac{2V_{TH}}{V_{amp}}\right) t_{rise}. \quad (3.3)$$

Regardless of the clocks and process, the period of this mode will be short relative to the reset and evaluation periods. The effective resistance of M_6 would only come into affect if the period was long and the capacitors had time to reach their equilibrium voltage of $0V$. Given this, we can model this mode as seen in Fig. 3.3 (b), which behaves the same as (a) at early times. This is especially true if the latch is designed well and R_{DS,M_6} is much larger than R_{DS,M_5} . This will be discussed in further detail later in this Chapter. The voltage on C_P as it discharges is given by

$$V_{C_P}(t) = (V_{init} - V_{eq}) \left(e^{\frac{-t}{R_{DS,M_5} \frac{C_T C_P}{C_T + C_P}}} \right) + V_{eq} \quad (3.4)$$

where V_{eq} is the equilibrium voltage and is given by

$$V_{eq} = V_{DD} \left(\frac{C_P}{C_P + C_T} \right) \quad (3.5)$$

and V_{init} is the initial voltage on C_P and is

$$V_{init} = V_{CM} - V_{TH} + \frac{V_{in}}{2}. \quad (3.6)$$

Similarly the voltage on C_T as it charges is given by

$$V_{C_T}(t) = (V_{eq}) \left(1 - e^{\frac{-t}{R_{DS,M_5} \frac{C_T C_P}{C_T + C_P}}} \right). \quad (3.7)$$

This mode begins when the gate voltage reaches V_{TH} and finishes when the gate is $V_{p,clock} - V_{TH}$, at which point the CS latch enters into the evaluation mode. It should be noted that R_{DS,M_5} is time-varying and is not presumed to be a constant here. However, taking the integral of R_{DS,M_5} over the period of Mode 2 will yield a constant equivalent resistance that can be utilized.

Mode 3: Evaluation

Once we enter into evaluation mode, M_6 turns OFF and M_5 has stronger conduction. Furthermore, due to the decrease in the voltage at node P during Mode 2, M_1 and M_2 have begun conducting and their effective resistance has decreased. One transistor will be operating with moderate inversion while the other transistor will have a lower gate voltage (by $V_{in,p}$) and will be operating in weak inversion. This gives us the circuit model seen in Fig. 3.4.

As in Mode 2, the voltage on the output nodes at the end of evaluation can be predicted. Unlike in Mode 2, it is not a simple case of two capacitors coming to equilibrium across a resistor. We will begin by looking at the voltage on C_{D1} . To start we will ignore one output node, say C_{D2} , since M_2 will be operating in weak inversion and will have a large drain-source resistance. This results in little charge-sharing between C_{D2} and the other capacitors. This means at the start of the period,

Figure 3.4: Charge-steering latch model during evaluation.

C_{D1} , C_P and C_T are attempting to come to a temporary equilibrium. C_{D1} will be discharging while C_P and C_T will be charging. If the latch stayed in this configuration for a long period of time the voltage on these capacitors would be

$$V_{eq1} = V_{C_{D1}} \frac{C_{D1}}{C_{total}} + V_{C_P} \frac{C_P}{C_{total}} + V_{C_T} \frac{C_T}{C_{total}} \quad (3.8)$$

where $V_{C_{D1}}$, V_{C_P} and V_{C_T} are the voltages on the capacitors to start the evaluation period and C_{total} is the summation of C_{D1} , C_P and C_T . To start evaluation, $V_{C_{D1}}$ and $V_{C_{D2}}$ are given by V_{DD} , V_{C_P} is given by (3.4) at t_{mode2} and V_{C_T} is given by (3.7) at t_{mode2} .

However, once these three capacitors reach temporary equilibrium, C_{D2} will be discharging slowly and as such C_{D1} , C_P and C_T will be charged by it. If given enough time, all four capacitors would reach equilibrium and would have the voltage given by

$$V_{eq2} = V_{C_{D1}} \frac{C_{D1}}{C_{total}} + V_{C_{D2}} \frac{C_{D2}}{C_{total}} + V_{C_P} \frac{C_P}{C_{total}} + V_{C_T} \frac{C_T}{C_{total}} \quad (3.9)$$

where C_{total} is the summation of all the capacitors: C_{D1} , C_{D2} , C_P and C_T .

In an optimally designed latch we do not want all four capacitors to come to equilibrium, instead we would like to see the evaluation mode end concurrently with C_{D1} , C_P and C_T reaching equilibrium with as little discharge on C_{D2} as possible. Accordingly, the transistors in a CS latch should be sized in such a way that this behavior is observed. This maximizes the differential swing, while limiting the single-ended swing, which in turn limits the power consumption.

Figure 3.5: Charge-steering latch model transient operation during evaluation.

If C_{D1} and C_T reach equilibrium at the same time, we have ensured that C_{D1} has discharged as much as possible. Ensuring C_{D1} discharges as much as possible ensures we have maximized the differential swing since it discharges much faster than C_{D2} . The evaluation period should end at this point. If the evaluation period continues then the only significant voltage change will be on C_{D2} . This will decrease the differential swing and increase the power consumption. The transient behavior of the CS latch model Fig. 3.4 can be seen in Fig. 3.5. We would like the evaluation period to end at “Partial Equilibrium.”

In reality, the temporary equilibrium voltage on C_{D1} , C_P and C_T will fall approximately halfway between (3.8) and (3.9). The voltages at the end of evaluation period should therefore be

$$V_{C_{D1}} = V_{C_P} = V_{C_T} = \frac{V_{eq1} + V_{eq2}}{2}. \quad (3.10)$$

C_{D2} can be modeled as a discharging capacitor, coming to equilibrium with another capacitor given by the combination of C_{D1} , C_P and C_T . The voltage on this capacitor is given by

$$V_{C_{D2}}(t) = (V_{DD} - V_{eq2}) \left(e^{\frac{-t}{R_{DS,M2} \frac{(C_{D2})(C_{D1}+C_P+C_T)}{C_{D2}+C_{D1}+C_T+C_P}}} \right) + V_{eq2} \quad (3.11)$$

where the final voltage will occur when $t = t_{per}/2 - t_{mode2}$, at which point the latch enters into the mode during which all switches are closed for the second time. It has been shown that, by the start of evaluation, V_{C_p} can discharge down from (3.6) a relatively far amount and with a low input swing the gate voltage on M_2 could be relatively high. This would result in M_2 being in saturation for a short period of time to start evaluation, until the drain voltage falls far enough. However, it is desirable to avoid this operation to ensure the output swing is sufficient and unnecessary power is not wasted. Therefore, M_2 can be treated as a resistor in a well designed CS latch.

It should be noted that [19] suggested that node P rises enough during evaluation to turn both M_1 and M_2 off [19]. This behavior would not only be undesirable as “Partial Equilibrium” would not be reached and potential swing would be wasted, but this behavior is also not seen. Since node P discharges significantly during Mode 2, the voltage on this node is low to start evaluation and never rises enough to turn both transistors off. This is independent of common-mode voltage since node P and the gate of the input pair both follow the common-mode voltage. This will be demonstrated in simulations later in Ch. 4.

Mode 4: All Switches Closed Prior to Reset

Similar to Mode 2, there is a period of time between evaluation and reset where all switches are closed. For half of this period the four nodes will have very little change in voltage. During the second half of this mode, Nodes X , Y and P will be charged towards V_{DD} and node T will discharge towards ground. Therefore this mode can be looked at as an extension of reset, the time of extension given by half of (3.2) or (3.3). With all this said, the exact change in voltage on each node during this time is not crucial to know when designing a CS latch. In order to understand why, we must understand how the performance of a CS latch will now be quantified.

3.1.2 Performance of the Basic Charge-Steering Latch

As discussed previously, there are two key metrics when quantifying the performance of a CS latch: differential swing and power consumption. We can now accurately predict the swing and power consumption. The swing is given by

$$V_p = V_{C_{D2}} - V_{C_{D1}} \quad (3.12)$$

where $V_{C_{D2}}$ and $V_{C_{D1}}$ are the voltages at the end of evaluation given by (3.10) and (3.11) respectively. Current will primarily be consumed in charging the capacitors and is given by

$$I = \frac{C_{D1}\Delta V_{C_{D1}}}{T_b} + \frac{C_{D2}\Delta V_{C_{D2}}}{T_b} + \frac{C_P\Delta V_{C_P}}{T_b} \quad (3.13)$$

where $\Delta V_{C_{D1}}$ and $\Delta V_{C_{D2}}$ are given by V_{DD} minus the final voltage on each node, again given by (3.10) and (3.11). ΔV_{C_P} is given by (3.6) minus the final voltage on its node during evaluation, also given by (3.10). T_b is no longer the clock period, but is now the period of during which the capacitors are charging. As previously mentioned, the capacitors charge for half of Mode 4 and all of reset, therefore $T_b = t_{per}/2 - t_{mode4}/2$. The understanding of the CS latch derived from this section will now be used to guide the design of a CS latch.

3.2 Design of the Basic Charge-Steering Latch

When designing a CS latch we will assume that two parameters are given: the speed at which the CS latch should operate and the output capacitance, C_D . In this tutorial, the CS latch will be sized as to maximize swing-to-power consumption ratio for these given parameters. The purpose of Ch. 3.1 was to provide in depth insight into CS latches and what is happening on each node. While this understanding is useful, when designing a latch we can simplify some of the analysis. First off we will look at the swing. As mentioned previously the output swing will be somewhere between the voltage on three capacitors ($C_{D1}/C_P/C_T$) coming to equilibrium and the voltage on all four capacitors ($C_{D1}/C_{D2}/C_P/C_T$) coming to equilibrium. C_P cannot be predicted

here as it is dependent on the transistor sizing, thus an approximation must be made. The initial charge on C_{D1} and C_{D2} will dominate the equilibrium expressions, as such a simplified estimate of the swing on the desired output node is given by the two output nodes and the tail node coming to equilibrium

$$\Delta V_{C_{D1}} = \frac{C_T}{C_T + 2C_D} V_{DD}. \quad (3.14)$$

While C_{D1} and C_{D2} will not equally effect the output swing, treating them as such helps to account for the missing impact of C_P . In turn we can estimate the swing on the tail node as

$$\Delta V_{C_T} = \frac{2C_D}{2C_D + C_T} V_{DD}. \quad (3.15)$$

Both (3.14) and (3.15) will overestimate the swing on their nodes because they do not account for parasitics at the output nodes nor for the charge on C_P and C_T to start evaluation, but they will be sufficient for initial design.

Equation (3.14) suggests that increasing C_T will continually increase the output swing, but this is not true. In [19] it is suggested that C_T not exceed C_D , given the swing is said to follow (2.1). The linear range of (2.1) is only up to $C_T = C_D$, as seen in Fig. 3.6, after which increasing C_T provides little increase in swing. However, the swing is not accurately given by one output capacitor discharging to the tail capacitor as suggested by (2.1), but rather is the result of four capacitors charge-sharing as described previously. While this means the swing is lower than predicted by (2.1), the output voltage is linearly related to the tail capacitor for a longer range. This is shown in Fig. 3.6, where, by taking the derivative of (2.1) and (3.14) with respect to the C_T/C_D ratio, it can be seen that increasing C_T/C_D in (3.14) gives a larger change in V_{out} than in (2.1). If parasitics at the output node are also accounted for then the linear range is extended further. Though there is a small dependency on the speed of the latch and transistor sizing, the tail capacitor should be in the range of $C_D < C_T < 4C_D$. Beyond this point there is little increase in the output voltage regardless of the tail capacitor as in Fig. 3.6. A good starting point for the tail capacitor is given by

Figure 3.6: Latch predicted swing for a given tail capacitor to output capacitor ratio.

$$C_T = 2C_D. \quad (3.16)$$

With C_T known, M_6 can now be sized. M_6 should be as small as possible while still allowing C_T to completely discharge to ground. Complete discharge of C_T must occur within the reset period. Knowing 99% of the discharge has occurred when $t = 5RC$, we can predict the equivalent resistance that must be seen across M_6 as

$$R_{DS,M_6} = \frac{t_{model}}{5C_T}. \quad (3.17)$$

The resistance of M_6 will be decreasing exponentially throughout the reset mode since its drain voltage is dictated by the discharge of C_T . Since the capacitor discharge and resistance of M_6 are dependent on one another, it is difficult to predict the exact voltage on the tail node at a given time. However, a decent prediction can be made. It is possible to predict the voltages on M_6 , and in turn the equivalent resistance of M_6 , at the start of reset and at the end of reset. The resistance in time can then be predicted by

$$R(t) = R_1 e^{\left(\frac{\ln(R_2/R_1)}{t_{mode2}}\right)t}. \quad (3.18)$$

We can then determine the equivalent resistance as

$$\begin{aligned} R &= \frac{1}{t_{mode2}} \int_0^{t_{mode2}} R(t) dt \\ &= \frac{(R_1 - R_2)}{\ln\left(\frac{R_1}{R_2}\right)} \end{aligned} \quad (3.19)$$

where R_1 is the drain-source resistance of M_6 to start reset and R_2 is the resistance to end reset. Setting (3.18) and (3.19) equal we can determine the time at which the equivalent resistance, given by (3.19), will be seen in (3.18) as

$$t = \frac{(t_{mode2}) \left(\ln \left(\frac{\frac{R_1}{\ln(R_1/R_2)} - \frac{R_2}{\ln(R_1/R_2)}}{R_1} \right) \right)}{\ln(R_1/R_2)}. \quad (3.20)$$

This time ends up being somewhere between 25% and 50% of the overall reset period. Given the resistance change and capacitance discharge follow a similar trend, C_T must have discharged between 25% and 50% when the resistance given by (3.17) is seen. As an initial assumption for the design equations it can be said that the resistance given by (3.17) should be seen when C_T has discharged one-third. At this point the gate voltage on M_6 will be halfway between $V_{clock,p} - V_{TH}$ and $V_{clock,p}$. The source is tied to ground and is 0V throughout. The drain voltage on M_6 will be low, limited by (3.15). Therefore, it will be conducting in the linear region of operation for nearly all of reset and as such we can predict the current across it as

$$I_{DS} = \mu_n C_{ox} \frac{W}{L} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]. \quad (3.21)$$

With sinusoidal clocks (dictating the gate voltage on M_6 is not high throughout reset) and a small tail capacitor (meaning less energy is lost in the transfer of charge from

C_D to C_T), it is possible that M_6 could be in saturation to start reset. However, the use of a small tail capacitor necessitates low speeds to ensure a sufficient output swing and at these speeds square clocks are more likely. Furthermore, by designing for triode conditions - which are much more likely - it is ensured that C_T will fully discharge. Knowing the voltages on each node and using (3.15), (3.17) and (3.18), the starting size for M_6 is given by

$$W_{M_6} = \frac{\frac{2\Delta V_{C_T} L}{3}}{R_{DS,M_6} \mu_n C_{ox} \left[(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]}. \quad (3.22)$$

It is impossible to treat M_5 in the same manner as M_6 since the capacitance at its drain is in part dependent on the size of M_5 , but we know several things about this transistor. Firstly, it is important that the resistance of M_5 is large so during Mode 2, C_P shares minimal charge with C_T . Secondly, as previously mentioned, we also want to ensure no energy is lost through M_6 's path to ground during Mode 2. The resistance of M_6 is as large as it can get while still ensuring C_T completely discharges, therefore, a large resistance across M_5 is the only other way to minimize energy loss during Mode 2. Thus, by making M_5 small we can maximize the equilibrium voltage during Mode 3 and in turn the output swing. A good starting point for M_5 is to make it half the size of M_6 . As the common-mode voltage increases, C_P will be sharing more charge with C_T during Mode 2, so there is the potential that more charge can be lost. As such the width of M_5 should have an inverse correlation with V_{CM} as well, such that its resistance is directly proportional to V_{CM} . If M_5 is made too much smaller we limit charge-sharing between C_D and C_T during Mode 3. These thoughts bring us to the following starting size for M_5

$$W_{M_5} = \frac{W_{M_6}}{2V_{CM}}. \quad (3.23)$$

M_3 and M_4 can be treated similar to M_6 . We know that these transistors must be sized in such a way as to allow the output nodes to charge back up to V_{DD} . Both nodes will discharge and need to be charged during reset, however the critical node will be the one that has discharged the most. Again, we can assume that 99% of the

charge will be delivered by $t = 5RC$. Therefore, the equivalent resistance of M_3/M_4 is given by

$$R_{DS,M_{3,4}} = \frac{t_{mode1}}{5C_D}. \quad (3.24)$$

Again, it can be said that we want to see this resistance when the drain is one-third charged ($V_{DD} - 2\Delta V_{C_D}/3$). The gate voltage will be at $V_{TH}/2$ at this point. The source voltage will be at V_{DD} throughout reset. From this we can size M_3/M_4 as

$$W_{M_{3,4}} = \frac{\left(V_{DD} - \frac{2\Delta V_{C_D}}{3}\right)L}{R_{DS,M_{3,4}}\mu_p C_{ox} \left[(V_{SG} - V_T)V_{SD} - \frac{V_{SD}^2}{2}\right]}. \quad (3.25)$$

By determining $R_{DS,M_{3,4}}$ using $t = 5RC$ it has been maximized and $M_{3,4}$ have been sized to the absolute minimum they can be. In some cases this means the output nodes may not fully charge to V_{DD} during reset, leading to ISI (also known as hysteresis in comparators) [19]. This design equation may yield a small amount of ISI, but if none can be tolerated than the width of $M_{3,4}$ can be increased until the differential output completely returns to zero during reset.

The sizing of M_1/M_2 is the final step in designing a CS latch. Every evaluation period we want one of these transistors to be small to create a large drain-source resistance and create a slow discharge. We want the other transistor to be large so the desired output node discharges quickly. The gate voltages of M_1/M_2 will differ by $V_{in,p}$ so one transistor will be operating in moderate inversion and the other will be in weak inversion, thus enabling the slow discharge across one input transistor and the quick discharge across the other input transistor. We can again give a good starting point for the input transistor in a manner similar to $M_3/M_4/M_6$.

We would like C_T and one C_D to come to equilibrium at the end of evaluation. In order for this to occur, the combined resistance seen across M_1 or M_2 and M_5 must be sized appropriately. While the drain of M_5 may charge a small amount during evaluation, the source will charge significantly, as denoted by (3.15). As such, M_5 should begin the evaluation period in saturation but as the tail node charges the transistor will enter triode and even subthreshold with sinusoidal clocks. This will

prevent the desired output node and the tail from ever actually reaching equilibrium. Instead it can be said that C_P and the desired C_D must reach equilibrium in an optimally designed latch. This means the resistance that must be seen across M_1 or M_2 is given by

$$R_{DS,M_{1,2}} = \frac{t_{mode3}}{5 \frac{C_D C_P}{C_D + C_P}}. \quad (3.26)$$

C_P and the width of the input pair depend on one another, thus making an accurate prediction at the parasitic capacitance difficult. But we can look at the relationship between C_D and C_P . It is known that the size of C_D is contributed to by the input pair, pMOS switches and the stage(s) which follow the latch. In an equalizer, the load capacitance will be significant due to the presence of taps and cascaded latches. In a PD, cascaded latches and XORs also result in a large C_D . C_P is only contributed to by both input transistors and the transistor M_5 . Regardless of application, C_P should be much smaller than C_D . It is not likely that C_P approaches even half of C_D , but we will design for this scenario to be safe. The equivalent resistance of $M_{1,2}$ can now be given as

$$R_{DS,M_{1,2}} = \frac{t_{mode3}}{\frac{5C_D}{3}}. \quad (3.27)$$

Again this resistance needs to be seen when the drain of M_1/M_2 has discharged a third (given by $V_{DD} - \Delta V_{C_D}/3$). The gate voltage will be $V_{CM} + V_{in,p}/2$ throughout the evaluation period and the source voltage should remain around $V_{CM} - V_{TH}$. For most of evaluation the desired transistor will be operating with moderate inversion. Using the EKV model for a transistor and assuming a slope factor of approximately 1, we can predict the current through a transistor operating in this region as [24]

$$I_{DS} = 2\mu C_{ox} V_T^2 \frac{W}{L} \left(\ln^2 \left(1 + e^{\left(\frac{V_G - V_{TH} - V_S}{2V_T} \right)} \right) - \ln^2 \left(1 + e^{\left(\frac{V_G - V_{TH} - V_D}{2V_T} \right)} \right) \right) \quad (3.28)$$

where V_T is the thermodynamic voltage, V_{TH} is the process threshold voltage and V_G , V_S and V_D are the gate, source and drain voltages, respectively. The desired width of the input transistors can now be determined by

$$W_{M_{1,2}} = \frac{(V_{DD} - \frac{\Delta V_{CD}}{3} - V_{CM} + V_{TH})L}{x_n R_{DS,M_{1,2}}} \quad (3.29)$$

where x_n is given by

$$x_n = 2\mu_n C_{ox} V_T^2 \left(\ln^2 \left(1 + e^{\left(\frac{V_G - V_{TH} - V_S}{2V_T} \right)} \right) - \ln^2 \left(1 + e^{\left(\frac{V_G - V_{TH} - V_D}{2V_T} \right)} \right) \right). \quad (3.30)$$

Equation (3.29) will set the width the input pair should be to achieve the optimal swing-to-power consumption ratio. Increasing the size of the input pair beyond (3.29) should not significantly affect the swing-to-power consumption because the desired output node will simply reach equilibrium early. The non-desired output node will discharge minimally for an increase in the input pair because the input transistor it discharges across should be in subthreshold regardless of size. With that said, sizing the input pair minimally limits the load capacitance seen by the previous circuit and should therefore be favored. It may even be beneficial to decrease the transistor from (3.29) if the loading of the previous stage is a limiting factor in design.

Clock skew has not been considered in the design equations above but this is appropriate for CS latches. The most critical switches (M_3/M_4 in reset and M_5 in evaluation) are driven by the same clock and, given the compactness of CS latches, there should be negligible skew between these switches. M_6 is being driven by the opposite clock, so skew would result in M_6 being closed during evaluation for a short period and M_5 being closed during reset for a short period. While this would result in a small amount of charge being wasted, for any reasonable skew the increase in power would be negligible. Even if M_6 was driven by the same clock as $M_3/M_4/M_5$, the circuit would still work, though with an increased power consumption.

These design equations will be verified later on but first other CS latches will be looked at in order to understand how to adapt this analysis for various applications.

3.3 Charge-Steering with pMOS Cross-Coupled Pair

As mentioned previously, adding a CC pair to basic CS latch - as in Fig. 2.13 - can help to pull one node back towards V_{DD} during evaluation. This can help to decrease the single-ended swing, thus, during reset, this node does not need to charge as much and power is saved. But these power savings are almost entirely canceled out by the current consumption of the CC pair, both during reset and in charging one node during evaluation. The current consumption of the CS latch with CC pair can be predicted by

$$I = \frac{C_D x \Delta V}{T_b} + I_{CC} \quad (3.31)$$

where x is a scaling factor to account for the drop in single-ended voltage swing and I_{CC} is the current consumption of the CC pair.

In order to ensure the CC pair functions as expected and charges only one output node during evaluation, M_7/M_8 must be sized appropriately. In order to achieve this behavior, the path from the non-desired node, for example Y , to the voltage supply must be less resistive than the path from Y to node P . As such, the resistance of M_8 should be less than the resistance of M_2 throughout evaluation. Conversely, we would like the path from X to the supply to be highly resistive to ensure no charging towards V_{DD} . As such, M_7 should be designed to conduct in the subthreshold region of operation. To ensure M_7 remains in subthreshold $V_{SG} - V_{TH} < -50mV$ [25], meaning the voltage on Y (the drain of M_2) cannot fall below $V_{DD} - V_{TH} + 50mV$ during evaluation. As previously discussed, the source voltage on M_2 will be around $V_{CM} - V_{TH}$ and the gate voltage will be $V_{CM} - V_{in,p}/2$. At the end of evaluation M_2 will still be in subthreshold and its current will be given by (3.28). With the current through M_2 and the voltages on its nodes known, we can now predict its equivalent resistance at the end of evaluation as

$$R_{DS1,2} = \frac{V_{DD} + 0.05 - V_{CM}}{I_{DS}}. \quad (3.32)$$

R_{DS,M_8} must at most be equivalent to R_{DS,M_2} . The source of M_8 will be at V_{DD} ,

the drain should be around $V_{DD} - V_{TH} + 50mV$ and the gate will be at the supply voltage minus the swing given by (3.14). With these voltages M_8 will be operating in triode and its minimum width to ensure charging will be given by

$$W_{M_{7,8}} = \frac{(V_{TH} - 0.05)L}{R_{DS,M_{1,2}}\mu_p C_{ox} \left[(V_{SG} - V_T)V_{SD} - \frac{V_{SD}^2}{2} \right]}. \quad (3.33)$$

This tells us the minimum width M_8 must be to ensure node Y charges more than it discharges, however M_8 should be as large as possible. It is only limited in size by making M_7 too large and causing it to operate outside subthreshold. We want the resistance across M_7 to be an order of magnitude larger than M_8 and the voltages on its nodes are the same as M_8 , except the drain and gate have swapped. The maximum width of M_7 can now be predicted as

$$W_{M_{7,8}} = \frac{\Delta V_{C_D} L}{x_p 10 R_{DS,M_{1,2}}} \quad (3.34)$$

where x_p is given by

$$x_p = 2\mu_p C_{ox} V_T^2 \left(\ln^2 \left(1 + e^{\left(\frac{V_S - V_{TH} - V_G}{2V_T} \right)} \right) - \ln^2 \left(1 + e^{\left(\frac{V_D - V_{TH} - V_G}{2V_T} \right)} \right) \right). \quad (3.35)$$

The optimal size of $M_{7,8}$ will tend towards (3.34), however, sweeping in the range given by (3.33) and (3.34) will allow the designer to maximize the swing-to-power consumption.

Figure 3.7: Charge-steering latch with clocked pMOS cross-coupled pair.

3.4 Charge-Steering with Clocked Cross-Coupled Pair

The pMOS CC pair is only required during evaluation mode, not when C_D is charging. As such it is desirable to effectively remove the pMOS pair from the circuit during reset. This thought leads to a new latch topology, which has a switch (M_9) in the CC pair as seen in Fig. 3.7.

In Fig. 3.8, the transient operation of the output nodes and the current consumption for both the basic CS latch and CS latch with clocked CC pair is shown. The discharge on the desired node follows a similar pattern regardless if a CC pair is present or not. However, in the basic latch, the non-desired node begins charging only once evaluation has ended, whereas with the CC pair, charging begins during evaluation. As a result there is a small current consumption during evaluation when the CC pair is present but there are current savings during reset, resulting in similar overall current consumption but higher output swing. The addition of the switch, M_9 , still allows for the charging of the non-desired node during evaluation but does not consume as much power during reset. The switch, therefore, allows for the CC pair to work effectively while reducing the overall power consumption when compared to both the basic latch and the latch with pMOS CC pair.

Figure 3.8: Charge-steering latch comparison.

Ch. 3.1.2 presented a more accurate prediction of the current consumption of the basic latch, but (2.2) and (3.31) will be used here to demonstrate the expected power savings. Presuming the CC pair current consumption is halved with the addition of this switch, the current consumption of this new topology is given by

$$I = \frac{C_D x \Delta V}{T_b} + \frac{I_{CC}}{2}. \quad (3.36)$$

The current consumption of the new CS latch can then be multiplied by a new scaling factor, y , and equated to (3.31). Solving this gives us the new scaling factor as

$$y = \frac{1 + \frac{C_D x \Delta V}{T_b I_{CC}}}{0.5 + \frac{C_D x \Delta V}{T_b I_{CC}}}. \quad (3.37)$$

Equations (2.2) and (3.31) could also be equated to solve for x and get (3.37) entirely in terms of C_D , ΔV , T_b and I_{CC} . The expected power consumptions savings of the new latch are correlated to y , which will be verified through simulation in Ch. 4. But first M_9 must be sized appropriately to ensure it works as expected.

During evaluation we would like this switch to have a small resistance so it does not interfere with the charging of Y . In order to achieve this it must be sufficiently larger than the pMOS CC pair. However, it should not be made too large that the clock buffers, which drive the switch, require more power than is saved in the actual latch. A good compromise is setting the width to twice that of the pMOS CC pair. The extra resistance added by the switch may necessitate the pMOS CC pair size is increased slightly from (3.33) and (3.34), but they still provide a good starting point.

In addition to the switch sizes changing, the tail capacitor can also be increased before hitting a saturation point for the swing-to-power consumption ratio. The linear range of the swing has been extended and the new transistors increase the parasitics at the output nodes such that the tail capacitor should now be tends towards $4C_D$.

The addition of another switch is of no concern as it pertains to voltage headroom. Not all transistors are closed at the same time and given the reset state of the latch and properly sized transistors, all nodes should reach their desired voltage by the time evaluation begins. Once evaluation begins, there are effectively only two stacked

transistors. This suggests that the new topology will also be applicable to future technology nodes.

While in depth analysis of the high-swing CS latches is outside the scope of this work, the insight provided herein can certainly be used to design the high-swing latches.

3.5 High-Swing Latches with Clocked Cross-Coupled Pair

Following the same logic as in Ch. 3.4, a switch (M_{18}) is added to pMOS CC pair in the high-swing latches. The new topologies can be seen in Fig. 3.9 and Fig. 3.10. This switch will again minimize the power consumption of M_7 and M_8 during reset while still benefiting from their presence during evaluation. Ultimately the same swing can be expected from these new topologies - when compared to the conventional high-swing latches - but at a lower power consumption.

Figure 3.9: High-swing charge-steering latch with clocked pMOS cross-coupled pair.

Figure 3.10: High-swing charge-steering latch with clocked pMOS cross-coupled pair and cross-coupled cascode devices.

3.6 High-Swing Latches with No Secondary Tail Capacitor

Adding the secondary tail capacitor to the high-swing latch may not be necessary in some applications. An alternative is to tie the source of the nMOS CC pair (M_{12}/M_{13}) to the source of M_1/M_2 [20]. Another option is to eliminate the secondary capacitor and have M_{12}/M_{13} switch directly to ground through M_{14} , similar to the pMOS CC pair being switched to V_{DD} . The new topologies can be seen in Fig. 3.11 and Fig. 3.12. The elimination of the secondary tail capacitor can be beneficial with or without the additional switch from Ch. 3.5 but for the sake of brevity, the new topologies have only been shown with the switch present.

The addition of M_{12}/M_{13} is desirable to achieve higher swings but eliminating the secondary tail capacitor will still increase voltage at the drain of the differential pair at a lower power consumption and lesser overall complexity. The advantage of eliminating the secondary tail capacitor is expected to be prevalent at higher data rates. This is because the secondary tail capacitor limits the influence of the nMOS cross-coupled pair, M_{12}/M_{13} . But as speed increases, the influence is already limited and to maximize the gain of this pair the secondary tail capacitor must be dropped.

Dropping a switch from the latch will also result in a reduced load on the clock buffers, which will save further power.

Figure 3.11: High-swing charge-steering latch with no secondary tail capacitor.

3.7 Charge-Steering Latch with Output Switch

It was mentioned earlier that the sizing of M_3 and M_4 has the most impact on CS latch performance. It was also mentioned that M_3/M_4 must be sized such that the output capacitor, which has discharged the most, is charged back up to V_{DD} . In this case one of M_3 or M_4 will be over-sized and charge will be wasted, but this is necessary to ensure an RZ output is seen. The foregoing thought leads to a new CS latch topology seen in Fig. 3.13. It can be seen that a small pMOS switch has been added between the basic CS latch's output nodes. Balancing two signals is common in comparator design, for example the original StrongARM latch [26], but has not been seen in CS circuits until now. This switch allows the output capacitor that reaches V_{DD} first to share some of its charge with the other capacitor such that no charge is wasted. The width of M_3 and M_4 can now be made smaller - which results in significant power savings in the latch itself - while still achieving an RZ output.

Figure 3.12: High-swing charge-steering latch with no secondary tail capacitor and cross-coupled cascode devices.

In addition to saving power in the latch, the load on the clock buffers will also be reduced. Intuitively, adding another switch would add load to the clock buffers, but this output switch does not need to be large, while the reduction in size of M_3 and M_4 will be significant.

There is one drawback to this topology. The design and layout are complicated slightly by having another transistor to size and more clock routing to perform. However, this drawback is minor and the power savings are almost certainly worth it. The output switch can be added to any of the aforementioned topologies while seeing the benefits.

Chapter 3.1 and 3.3 analyzed the basic CS latch and CS latch with a CC pair in a novel way. This analysis was used to develop a design methodology for these latches and also led to the development of several new topologies. The new topologies added a switch to the CC pair as well as the between the output nodes and removed the secondary tail capacitor from the high-swing latches, leading to an improved performance. But in order to confirm the efficacy of the design equations and new topologies from this Chapter, simulations are required.

Figure 3.13: Basic charge-steering latch with output switch.

Chapter 4

Simulation of Charge-Steering Latches

Chapter 4.1 simulates the basic CS latch in two CMOS processes. First off simulations are performed in STMicroelectronics' 28nm Ultra Thin Body and Buried Oxide Fully Depleted Silicon on Insulator Complementary Metal Oxide Semiconductor (UTBB FD-SOI CMOS) process, which uses a Leti-UTSOI2 transistor model. Following this, simulations are performed in TSMC's 65nm bulk CMOS process. In both processes, simulations are performed to compare the design equations for the basic CS latch to the optimal point of operation in the given process. In Ch. 4.2, 4.3 and 4.4, the new CS latch topologies are simulated and compared to simulations for the conventional CS latches as well as the StrongARM latch. Throughout this Chapter, in all simulations, the latches will be tested with random input data with a swing of 100mVp and NRZ encoding. All clocks will have an amplitude of 1V and will be at a frequency equal to the data rate. The output capacitance, C_D , is always 10fF and the supply voltage is always 1V.

Prior to these simulations, some earlier claims must first be verified. The worst case scenario for the latch must be simulated to verify the claim that node P does not rise enough during evaluation to turn off both M_1 and M_2 . To verify this claim the data rate was set high (28Gb/s) so that the period of Mode 2 is short, and therefore node P cannot discharge for long during Mode 2. To further shorten the period of Mode 2, square clocks were used with fast rise/fall times (5%). Furthermore, a low common-mode voltage (500mV) and input swing (100mVp) were used to ensure the gates voltage on M_1 and M_2 were low to start evaluation. M_5 was sized small ($2\mu\text{m}$ width), at the cost of swing, to ensure it has a high resistance during Mode 2. Finally, the smallest reasonably sized tail capacitor (1:1 ratio with the output capacitor) was used to ensure it attracted as little voltage from node P as possible. Other transistors

were sized rather generically, the input pair width to $3\mu\text{m}$ and the pMOS switches to $4\mu\text{m}$. Lengths were set to the process minimum.

The results of this initial simulation can be seen in Fig. 4.1. It can be seen that, even in this worst case scenario, the voltage on node P drops significantly during Mode 2 down to 130mV. While it does rise slightly during evaluation (to 180mV), V_{GS} never dips below the processes threshold voltage and discharge of the output capacitors never completely ceases. With a low common-mode voltage, it is also not possible for V_{DS} to dip below V_{GS} at any point as the output capacitors will be discharging relatively slowly. Any increase in common-mode voltage will allow the output nodes to discharge further, but would equally raise the gate voltage and node P such that triode or weak inversion conditions will never be satisfied for one of the input transistors. If M_1 never completely turns off, the discharge of its capacitor will be ceased by reaching “Partial Equilibrium”, at which point C_{D_2} will continue to discharge - decreasing the differential swing while simultaneously wasting charge - should the evaluation mode not be ended. As such, the previous stated goal of the design equations (for C_P and one of the output capacitors to reach “Partial Equilibrium”) has been verified as resulting in the optimal design.

Another claim that must be verified via simulation, is that Fig. 3.3 (a) and (b) act the same for short periods. To verify this, the simulation set ups in Fig. 4.2 (a) and (b) were used. The initial voltages on C_P and C_T were set to 1V and 0V, respectively, for both models. In (a), the gate of M_5 was ramped up from V_{TH} to $V_{amp} - V_{TH}$ and the gate of M_6 from $V_{amp} - V_{TH}$ to V_{TH} , as occurs during Mode 2. The period of the voltage ramp was varied from 16.09ps to 2.88ps to reflect the period of Mode 2 for speeds of 5Gb/s to 28Gb/s in the 28nm process with sinusoidal clocks. In (b), M_6 was dropped to demonstrate that it has little effect for small periods. C_P was sized to 10fF and C_T to 20fF. All transistors, in both simulations, were sized to 2μ . The voltage on C_P for each of these simulations can be seen in Fig. 4.3. Regardless of speed it can be seen that the voltage on C_P in both models is nearly identical until well past the period of Mode 2, verifying the claim that Fig. 3.3 (a) and (b) act the same for short periods, in situations realistic to CS latches.

Figure 4.1: Basic charge-steering latch transient operation in worst case scenario, demonstrating one of the input differential pair transistors never turns off during evaluation.

Figure 4.2: Basic charge-steering latch equivalent model during Mode 2. a) Full model. b) Simplified model.

A final claim that must be verified is that CS circuits are not susceptible to clock skew. To verify this claim, simulations were performed on the basic CS latch in 28nm at 10Gb/s with random NRZ data that had a 100mVp input swing, 850mV common-mode voltage and 1Vpp sinusoidal clocks. Transistors and capacitors were sized generically - input pair width set to $4\mu\text{m}$, switches to $3\mu\text{m}$, load to 10fF and tail capacitor to 20fF. Skew was then added between the two clocks ranging from 0% to 100%. The resulting swing and power consumption are plotted in Fig. 4.4.

Since M_6 is closed for a period of evaluation when it should be open, some power is wasted as expected, but even at 100% skew (M_6 closed for all of evaluation) the circuit still operates. The swing is reduced by 8.1% with 100% skew and the power consumption a little more than doubles. However, for any reasonable skew (up to 10%), the increase in power consumption is small (less than 7%), while the swing remains nearly identical. This confirms that clock skew is not of primary concern when designing CS latches.

Figure 4.3: Voltage on node P during Mode 2 of basic charge-steering latch model.
 (a) 5 Gb/s. (b) 10 Gb/s. (c) 14 Gb/s. (d) 20 Gb/s. (e) 28 Gb/s.

Figure 4.4: Clock skew effect on basic charge-steering latch swing and power consumption.

4.1 Simulation of Basic Charge-Steering Latch

To verify the design equations for the basic latch provided in Ch. 3, this section will present simulations in two different processes. Simulations were performed in STMs' 28nm process and in TSMC's 65nm bulk CMOS technology.

In 28nm three distinct simulation types were performed at data rates of 5, 10, 14, 20 and 28Gb/s. First off, simulations were performed utilizing the design equations to size the transistors. Second, a larger simulation was performed where each transistor size was swept in order to determine the actual optimal point of the CS latch. The purpose of the sweeping simulation is to demonstrate how close the design equations get us to the optimal point with minimal work. Following these simulations layout was performed to verify that the design equations hold up when extracted.

In 65nm, the basic CS latch was also simulated at data rates of 5, 10, 14, 20 and 28Gb/s. Simulations were once again performed using a sweeping simulation and the design equations. Simulations of CS latches using sizes from [13] and [19], which both utilize the same process, were also performed. The focus of [13] and [19] is on larger systems and they do not claim optimally sized CS latches, but these simulations provide insight into the type of improvement the design equations can provide. All simulations which follow use sinusoidal clocks, to replicate the worst case scenario and the common-mode voltage is 850mV.

Table 4.1: Basic latch component sizes for sweeping simulations in 28nm

Component	Minimum Size	Maximum Size	Step Size
M_3/M_4	$0.5\mu\text{m}$	$10\mu\text{m}$	$0.5\mu\text{m}$
M_5	$0.25\mu\text{m}$	$2.5\mu\text{m}$	$0.25\mu\text{m}$
M_6	$0.25\mu\text{m}$	$5\mu\text{m}$	$0.25\mu\text{m}$
C_T	10fF	40fF	5fF

4.1.1 Basic Charge-Steering Latch 28nm Simulations

To begin simulating the basic CS latch, the input differential pair was sized at $5\mu\text{m}$. As mentioned in Ch. 3.2, the size of the input pair is not as crucial as the other transistors and $5\mu\text{m}$ was found to be an appropriate size regardless of data rate. Table 4.1 summarizes the remaining components and the values for which they were swept.

In order to determine the optimal point, a Figure of Merit (FoM) was defined as

$$FoM = \frac{V_{out,peak}}{Power\ Consumption}. \quad (4.1)$$

The output was monitored in each simulation ensuring it was RZ. As mentioned in Ch. 3, by sizing $M_{3,4}$ small, the FoM will significantly improve though at the cost of ISI. Such designs were discounted in the following simulations and the output was never more than $\pm 25\text{mV}$ during the RZ state. For each data rate the optimal FoM was determined.

Following the sweeping simulations, the design equations (3.22), (3.23) and (3.25) were used leaving the tail capacitor at the size found in the sweeping simulation. Table 4.2 summarizes the component size for each data rate. Finally, using the sizes from Table 4.2 the basic CS latch was laid out to be DRC and LVS compliant. A sample layout can be seen in Fig. 4.5.

Fig. 4.6 shows the FoM for each of the three simulations described above. It can be seen that the layout leads to a degradation in performance of at most 7% verifying the accuracy of CS latch simulations. It can also be seen that the design equations

Table 4.2: Basic latch component sizes from design equations in 28nm

Component	5 Gb/s	10 Gb/s	14 Gb/s	20 Gb/s	28 Gb/s
M_3/M_4	$0.57\mu\text{m}$	$1\mu\text{m}$	$1.5\mu\text{m}$	$2\mu\text{m}$	$2.6\mu\text{m}$
M_5	$0.16\mu\text{m}$	$0.356\mu\text{m}$	$0.5\mu\text{m}$	$0.81\mu\text{m}$	$1.3\mu\text{m}$
M_6	$0.256\mu\text{m}$	$0.6\mu\text{m}$	$0.84\mu\text{m}$	$1.4\mu\text{m}$	$2.2\mu\text{m}$
C_T	25fF	25fF	25fF	25fF	25fF

Figure 4.5: Layout for basic charge-steering latch in 28nm CMOS.

are, in the worst case, within 7% of the sweeping simulations, and match perfectly at high-frequencies, demonstrating their accuracy.

Figure 4.6: Figure of merit comparison for basic latch in 28nm.

To verify equation (3.29) and the behaviour of the input differential pair as described in Ch. 3.2, the input pair width was also swept from $1\mu\text{m}$ to $12\mu\text{m}$ for each data rate. All other components were sized as in Table 4.2. Fig. 4.7 shows the results of this simulation. It can be seen that regardless of the input pair size the FoM stays relatively constant. At each data rate the input pair size recommended by (3.26) is marked. In each case sizing M_1/M_2 using (3.26) gives the highest possible FoM or very near that point.

Figure 4.7: Figure of merit for basic latch in 28nm for various input transistor widths.

A quick note on layout of the basic CS latch. The layout of individual CS latches is not particularly challenging. The main concern regarding the layout of these latches comes at the system level when there will need to be significant clock distribution. This is certainly a drawback of CS latches but the power savings can be worth the extra effort required in layout.

There some advantages to the layout of CS latches, in particular the area. To demonstrate where these advantages come from I will use an example. The layout of the basic CS latch at 5Gb/s from Fig. 4.5 occupies an area of $7.5\mu\text{m}$ by $3.4\mu\text{m}$. Using (2.4), we can estimate that a CML latch would require a resistor of 4500Ω at this data rate. In the 28nm FD-SOI process, an n-well resistor of this size would occupy a space of $10\mu\text{m}$ by $2.1\mu\text{m}$. This tells us that one of the required resistors for a CML latch nearly matches the layout area of the entire CS latch. Accounting for both resistors and all other devices suggests that a CML latch would occupy approximately twice the area of a CS latch.

Table 4.3: Basic latch component sizes for sweeping simulations in 65nm

Component	Minimum Size	Maximum Size	Step Size
M_3/M_4	$1\mu\text{m}$	$10\mu\text{m}$	$1\mu\text{m}$
M_5	$0.25\mu\text{m}$	$2.5\mu\text{m}$	$0.25\mu\text{m}$
M_6	$0.5\mu\text{m}$	$5\mu\text{m}$	$0.5\mu\text{m}$
C_T	10fF	40fF	5fF

Table 4.4: Basic latch component sizes from design equations in 65nm

Component	5 Gb/s	10 Gb/s	14 Gb/s	20 Gb/s	28 Gb/s
M_3/M_4	$1.5\mu\text{m}$	$3\mu\text{m}$	$4.2\mu\text{m}$	$6\mu\text{m}$	$8.4\mu\text{m}$
M_5	$0.55\mu\text{m}$	$1.1\mu\text{m}$	$1.54\mu\text{m}$	$2.2\mu\text{m}$	$3.07\mu\text{m}$
M_6	$0.933\mu\text{m}$	$1.87\mu\text{m}$	$2.61\mu\text{m}$	$3.73\mu\text{m}$	$5.23\mu\text{m}$
C_T	25fF	25fF	25fF	25fF	25fF

4.1.2 Basic Charge-Steering Latch 65nm Simulations

The next simulations were performed in TSMC's 65nm Bulk CMOS process. Again a sweeping simulation was performed for the basic latch and the optimal sizes were determined. The input differential pair size was reduced to $4\mu\text{m}$. Remaining sizes are given in Table 4.3. The design equations were also again applied. The sizes can be seen in Table 4.4.

The optimal FoM found from the sweeping simulations and the FoM found using the design equations can be seen in Fig. 4.8. The design equations yield a FoM that is within 11% of the FoM given by sweeping the transistors. This demonstrates the ability of the design equations from Ch. 3 to work in different processes.

To compare the design equations to prior art, first the basic CS latch from [19] was simulated. Measurement results for the basic CS latch are given in [19] in TSMC's 65nm process. All input parameters and sizes are given in addition to the transient operation of the latch on each node. These results were replicated in simulation to ensure a one-to-one comparison with prior art.

Figure 4.8: Figure of merit comparison for basic latch in 65nm.

Once the results were replicated, the latch was simulated using the input parameters from previous simulations (common-mode of 850mV, input swing of 100mV, 1Vpp sinusoidal clocks, random NRZ data) but component sizing from [19]. In addition to this, in [13] a CS flipflop is presented, a component of which is a CS latch and again all sizes are known. This latch was also simulated using these simulations parameters. Following this, for the same speed and input parameters, the basic CS latch was simulated using the design equations. Table 4.5 shows the transistor sizes for these two papers along with those determined by the design equations for multiple speeds. Also in the Table 4.5 is the swing, power consumption and FoM found for each latch. As seen in Table 4.5 and as highlighted in Fig. 4.9, the design equations yield FoM improvements of 350% and 185% at 2Gb/s and 12.5Gb/s respectively.

Figure 4.9: Performance comparison with prior art in 65nm.

Table 4.5: Basic latch component sizes comparison in 65nm

Reference	[19]	This Work @ 2Gb/s	[13]	This Work @ 12.5Gb/s
Data Rate	2Gb/s	2Gb/s	12.5Gb/s	12.5Gb/s
M_1/M_2	$4\mu\text{m}$	$4\mu\text{m}$	$4\mu\text{m}$	$4\mu\text{m}$
M_3/M_4	$8\mu\text{m}$	$0.6\mu\text{m}$	$8\mu\text{m}$	$3.75\mu\text{m}$
M_5	$8\mu\text{m}$	$0.6\mu\text{m}^*$	$8\mu\text{m}$	$1.37\mu\text{m}$
M_6	$4\mu\text{m}$	$0.6\mu\text{m}^*$	$4\mu\text{m}$	$2.33\mu\text{m}$
C_T	5fF	25fF	20fF	25fF
Swing	278mV	392mV	257mV	249mV
Power Consumption	$120\mu\text{W}$	$48\mu\text{W}$	$672\mu\text{W}$	$351\mu\text{W}$
Figure of Merit	2321	8106	383	709

*Calculations yielded a width smaller than process minimum so width was set to process minimum.

4.2 Comparison of Low-Power Charge-Steering Latches

The design equations presented in this thesis for the basic CS latch demonstrate the ability to significantly improve the latch performance. However, the CS latch with a pMOS CC pair and the new latch with a clocked CC pair are expected to further affect performance. This section compares these three latches to demonstrate how much an improvement can be expected from the new low-power latch topology.

The design equations were not used here, instead the transistors were sized rather generically in order to ensure the latches work at all data rates (5Gb/s to 28Gb/s). In this section, for all the latches, the differential pair size was set to $3\mu\text{m}$. For the basic latch the pMOS and nMOS switches were sized to $4\mu\text{m}$ and $3\mu\text{m}$ respectively. The width of the pMOS CC pair was set to $1\mu\text{m}$, keeping all other transistors the same size. For the latch with the clocked CC pair, the new switch was twice as wide as the other pMOS switches, $8\mu\text{m}$.

There are two points of performance which will be focused on for all latches in this section: the output amplitude and the power consumption. Transient operation

of a basic CS latch can be seen in Fig. 4.10. It can be seen that the output is RZ and the amplitude is not constant. Only the critical amplitude, defined as the amplitude following several identical digits, will be considered here.

Figure 4.10: Basic charge-steering latch transient operation.

Regardless of data rate, the new low-power latch reaches the same amplitude level as the other two low-power latches at a lesser power consumption as seen in Fig. 4.11. Corner simulations were also performed to determine whether the new latch achieves the power savings in the following five cases: fast nMOS-slow pMOS (FS), slow nMOS-fast pMOS (SF), fast nMOS-fast nMOS (FF), slow nMOS-slow pMOS (SS), typical nMOS-typical pMOS (TT). The performance of the new latch across all these corners can be seen in Fig. 4.12. The biggest degradation in performance comes from the FS simulations and the degradation in performance is only significant at lower data rates.

Fig. 4.13 plots the power consumption of the three low-power latches at the peak amplitude of the basic CS latch. The predicted power consumption of the new latch, given by (3.37), is also plotted. The data used in (3.37) was collected using the tail capacitor that corresponds to the peak amplitude of the basic latch. The predicted power consumption closely matches the simulated power consumption verifying the accuracy of (3.37). The slight differences between predicted and simulated power

consumption for the new latch are due to the predictions only accounting for charging the load capacitors and not the current consumed in charging parasitic capacitances.

Power savings were found using

$$Power\ Savings = 100 - 100 \frac{P_{clockedCC}}{P_{basic}} \quad (4.2)$$

where $P_{clockedCC}$ is power consumption of the new latch and P_{basic} is power consumption of the basic latch. The new latch topology achieves power savings of 10-40% depending on the data rate. For the FS process corner, the power consumption of the new latch at the peak amplitude of the basic CS latch is also plotted in Fig. 4.13. Even without accounting for the degradation in performance of the other two low-power latches due to process corners, the new latch still outperforms conventional CS latches with power savings of 10-35%.

Figure 4.11: Critical amplitude versus power consumption for low-power latches.
 (a) 5Gb/s data rate. (b) 10Gb/s. (c) 14Gb/s. (d) 20Gb/s. (e) 28Gb/s.

Figure 4.12: Corner simulations for charge-steering latch with a clocked cross-coupled pair. (a) 5Gb/s data rate. (b) 10Gb/s. (c) 14Gb/s. (d) 20Gb/s. (e) 28Gb/s.

4.2.1 Effect of Clock Buffers on Charge-Steering Latches

These power savings do not account for the additional load that the switch adds to the clock buffers. In order to assess whether the new topologies are still beneficial once clock buffers have been accounted for, a deeper understanding of SerDes design is required. As discussed previously, a full SerDes will see several lanes used in parallel. The most power hungry and accurate part of the clock generation will be shared between multiple SerDes. In recent designs one accurate Phase-Locked Loop (PLL) has been used and then distributed to Injection Locked Oscillators (ILOs) which are located in each lane of the SerDes [27], [28]. In this approach clock drivers are required which provide gain in the clocking path. This accounts for the amplitude that is lost as the clock signal travels across metal traces. Each switch will add load to the clock buffers due to the parasitic capacitance at the gate. The power consumption of a clock buffer to drive a single switch can be estimated by

$$P_{buffer} = CV_{amp}^2 f \quad (4.3)$$

where C is the parasitic capacitance seen at the gate of the switch, V_{amp} is the clock swing and f is the frequency of operation. In the previous simulations the new switch was sized to $8\mu\text{m}$, which has a measured gate capacitance of 5.5fF . This power has been added to the power of the CS latch with a clocked CC pair in Fig. 4.14. While the power savings have been lost at 28GHz , there are still savings of 5-30% at all

Figure 4.13: Power consumption at peak amplitude of low-power latches with comparison to predicted power savings.

other frequencies. Furthermore Fig. 4.14 and Fig. 4.13 are only comparing the power consumption of the latches at the peak amplitude of the basic latch. As such they do not represent the significant amplitude increase that can be gained by slightly increasing the power consumption of the CS latch with a clocked CC pair.

Figure 4.14: Power consumption at peak amplitude of low-power latches including power added by clock buffers.

In CS based transceivers the large number of switches may mean that having the clock generation circuitry repeated in each lane is beneficial from a power savings standpoint. This can help eliminate clock buffers all together. In [15] the parasitic capacitance of switches is absorbed into the VCO design and clock buffers are not utilized. Furthermore, extensive hardware sharing is done between the CDR and equalizers to minimize the additional power that having an individual CDR in each lane otherwise adds. While having clock circuitry in each lane would add to the die size, the power savings may make it worth it to do so. In such a design the power savings of the CS latch with a clocked CC pair are as in Ch. 4.2 and the additional switch's load need not be considered.

While the efficacy of the new CS latch with a clocked CC pair has been demonstrated, it is important that this latch is sized properly to get the power savings and swing improvement. The design equations for this latch from Ch. 3.2, 3.3 and 3.4 will now be verified in a manner similar to how they were confirmed for the basic CS latch in Ch. 4.1.1 and 4.1.2.

Table 4.6: Charge-steering latch with clocked cross-coupled pair component sizes for sweeping simulations in 28nm

Component	Minimum Size	Maximum Size	Step Size
M_7/M_8	$0.5\mu\text{m}$	$5\mu\text{m}$	$0.5\mu\text{m}$
M_9	$1\mu\text{m}$	$10\mu\text{m}$	$1\mu\text{m}$
C_T	10fF	50fF	5fF

4.2.2 Charge-Steering Latch with Clocked Cross-Coupled Pair 28nm Simulations

The same simulation setup that was used for the basic CS latch was also used to simulate the latch with a clocked CC pair in STMs' 28nm process. All transistors and the load capacitors from the basic CS latch were sized identically in the latch with a clocked CC pair. Then M_7 , M_8 , M_9 and the tail capacitor were swept in order to determine the highest possible FoM for each data rate. Table 4.6 summarizes the sizes for which simulations were performed. As before, the optimal FoM was determined for each data rate.

Following this (3.34) was used to size M_7 and M_8 . M_9 was sized to twice that of the cross-coupled pair. For all data rates the tail capacitor was left at 40fF, M_7 and M_8 were $2.4\mu\text{m}$ and M_9 was $4.8\mu\text{m}$. Their sizes did not need to scale with data rate since the input differential pair was left at $5\mu\text{m}$ for all data rates. Again, the CS latch was laid out using the sizes from the design equations and a sample can be seen in Fig. 4.15.

Fig. 4.16 summarizes the results of the sweeping, design equation and post layout simulations. As before, the layout leads to a small degradation in performance when compared to the schematic but the FoM is within 17% regardless of data rate. The FoM from the design equations are within 3% of the sweeping simulation regardless of data rate. This demonstrates the accuracy of the design equations and provides a foundation for designing further latches such as the high-swing latches, which will be discussed next.

Figure 4.15: Layout for charge-steering latch with clocked cross-coupled pair in 28nm CMOS.

Figure 4.16: Figure of merit comparison for charge-steering latch with clocked cross-coupled pair in 28nm.

4.3 Simulation of High-Swing Latches

The same simulation setup that was used for the low-power latches was also used to simulate the high-swing latches in STMs' 28nm process. Six high-swing latches were looked at: the two conventional high-swing latches, the two new high-swing latches with a switch added to pMOS CC pair and finally the two latches with no secondary tail capacitor. As was found for the low-power latches, adding a switch to the pMOS CC pair in the high-swing latch decreases the power consumption regardless of data rate as seen in Fig. 4.17. At lower data rates, it is preferable to have a tail capacitor; however, as data rates go to 14Gb/s and beyond, it is preferable to have no tail

capacitor. This is evidenced by Fig. 4.18, which plots the power consumption at which each latch achieves the peak amplitude of high-swing latch 1 from Fig. 2.14. In some cases it may even be beneficial to eliminate the primary tail capacitor and allow the parasitics to draw charge from the load capacitors. This is evidenced by the consistent amplitude of high-swing latch 1 with no secondary tail capacitor. As seen in Fig. 4.17 at 10Gb/s and beyond the amplitude for this latch is approximately the same regardless of the power consumption, which is directly correlated to the primary tail capacitor. While it appears that high-swing latch 1 and its variations are the best choice of topology, it should be noted that high-swing latch 2 and its variations can achieve a much higher peak swing at the cost of increased power consumption. The four new latch topologies achieve power savings of anywhere from 10-36% if the appropriate topology is chosen for each data rate, when compared to conventional high-swing CS latches.

CS latches are also compared to the StrongARM latch as seen in Fig. 2.12. The StrongARM is a more conventional high-swing, high-speed latch [23], [29]. The original StrongARM latch was slightly modified in my simulations. The tail transistor was not fed a clock signal, instead a DC voltage was used to create a constant current source in the tail. While this causes the StrongARM to consume static power it also extends the speed of operation, a desirable feature in transceiver latch design. The pMOS transistors and nMOS in the tail were sized to $6\mu\text{m}$, while the remaining nMOS transistors were sized to $3\mu\text{m}$. The DC voltage source was swept from 300mV to 700mV to mimic sweeping the capacitor in the tail of the CS latches. As seen in Fig. 4.18, regardless of data rate, all high-swing CS latches outperform the StrongARM when it comes to power consumption. The caveat to this is the StrongARM achieves rail-to-rail swing up to 9Gb/s. However, beyond this data rate the StrongARM's output amplitude decreases quickly and CS latches are preferable for high-swing as well. While the StrongARM is an excellent latch for lower speed applications, it cannot be the primary latch used in a high-speed, low-power transceiver, which is attempting to push 56Gb/s at a power efficiency of 1mW/Gb/s.

Figure 4.17: Critical amplitude versus power consumption for high-swing latches. (a) 5Gb/s data rate. (b) 10Gb/s. (c) 14Gb/s. (d) 20Gb/s. (e) 28Gb/s.

Figure 4.18: Power consumption at peak amplitude of high-swing latches.

4.4 Simulation of Charge-Steering Latch with Output Switch

To compare the basic CS latch to the new CS latch with an output switch, simulations were performed in 28nm FD-SOI. The widths of M_1 , M_2 , M_5 and M_6 were sized to $3\mu\text{m}$, for both latches. The tail capacitor was set to 20fF for a 10fF load. The common-mode voltage was set to 500mV. Remaining parameters were set as in previous simulations.

In Fig. 4.19, simulations are shown for the basic CS latch with the width of M_3/M_4 set to $0.5\mu\text{m}$ and $1.5\mu\text{m}$ at 10Gb/s. With M_3/M_4 set to $0.5\mu\text{m}$, the output does not return-to-zero, but the current consumption is low ($I_{AVG} = 62\mu\text{A}$). With M_3/M_4 set to $1.5\mu\text{m}$, the output does return-to-zero, but the current consumption is much higher ($I_{AVG} = 91\mu\text{A}$). Also shown in Fig. 4.19, is a simulation of the new CS latch with an output switch. M_3/M_4 and the output switch are all set to $0.5\mu\text{m}$. Despite the small switch sizes, the new latch's output is able to return-to-zero, while the power is nearly identical to the basic latch with no output switch and $0.5\mu\text{m}$ M_3/M_4 . The power savings of the new latch amount to 32%. This doesn't include the power savings in the clock drivers. The total width of M_3 , M_4 and the output switch has been reduced from $3\mu\text{m}$ in the basic latch to $1.5\mu\text{m}$ in the new latch, amounting to power savings of 50% in the clock buffers, only considering these switches.

The same simulations, described above for 10Gb/s, were also performed for data rates of 5Gb/s, 14Gb/s, 20Gb/s and 28Gb/s. The widths of M_3 and M_4 were swept at each speed until the minimum size, that caused the differential output to return-to-zero, was found. The output switch was sized to $0.5\mu\text{m}$ for all speeds. The results of these simulations are shown in Fig. 4.20, where the power consumption of the basic latch and new latch are plotted against data rate. It can be seen that regardless of data rate, the new latch achieves power savings of 23-33%. This result was found to be true across all process corners.

To give an idea of the parasitic capacitance savings that can be had by adding a switch between the output nodes, the total width of M_3/M_4 for the basic latch is compared to the total width of M_3/M_4 and the output switch for the new CS latch in Fig. 4.21. The total width of the three reset switches for the new latch range from 21% to 60% smaller than the total width of the two reset switch in the basic latch, data rate dependent. These width savings directly correspond to parasitic capacitance

savings and are the reason for the majority of the power savings in the new design.

To begin this Chapter, the basic CS latch was simulated in multiple processes; STMs' 28nm and TSMC's 65nm. In both processes it was found that the design equations could yield a latch whose FoM was close to the maximum possible FoM in a given process. This was confirmed with extracted simulations. Following this, Ch. 4.2 compared the basic CS latch, CS latch with CC pair and the new latch, with a switch added to the CC pair. Chapter 4.3 simulated and compared conventional high-swing latches to the new latches with no secondary tail capacitor and a switch in CC pair. In each case it was found that the new topologies could achieve power savings. Chapter 4.4 simulated and compared the new CS latch with an output switch to the basic CS latch, which again demonstrated power savings of as much as 33%. Overall, this Chapter demonstrates the efficacy of the new analysis, design equations and CS topologies presented in Ch. 3.

Figure 4.19: Basic charge-steering latch and charge-steering latch with output switch transient simulations at 10Gb/s.

Figure 4.20: Power consumption of basic charge-steering latch versus charge-steering latch with output switch.

Figure 4.21: Total width of all precharge switches in the basic charge-steering latch versus charge-steering latch with output switch.

Chapter 5

Conclusion

This Chapter summarizes the organization of this thesis and briefly comments on the contributions this thesis has made to wireline transceiver design and CS circuits in general. Chapter 5.1 then looks at the future work that could be done as an extension of this thesis.

This thesis began by exploring the use of CS latches in modern transceiver design. The basic CS latch is analyzed in a novel way yielding a new design methodology to help in sizing CS latches. This analysis is extended to several other conventional CS latch topologies. The insight learned from this is utilized to develop several new and improved CS latch topologies.

Simulations, both schematic and post-layout, are performed to demonstrate the validity of the new design equations and topologies. Regardless of the process, the design equations yield a latch whose FoM is within 11% of the highest possible FoM for the given process. When compared to existing CS latches in literature, the design equations yield improvements in the swing and power consumption. CS latches previously reported in literature have been used in systems, so the FoM improvement from the design equations only pertains to some of the circuits in these systems. Nonetheless, the overall system would see improvements in power consumption by using the design equations. Adding a switch to the pMOS CC pair allows the new topologies to achieve power consumption savings of 10-40%, data rate dependent. Removing the secondary tail cap was found to provide power savings of 10% at higher data rates. Adding a switch between the basic latch's output nodes was found to provide power savings as high as 33%.

In all, the improvement provided to CS circuits in this thesis can enable a 56Gb/s receiver with a 1mW/Gb/s power efficiency, or can further improve upon the power

consumption of the 40Gb/s, 14mW receiver presented in [16].

5.1 Future Works

While the new topologies and design methodology show promise there is still much work to be done in the study of CS latches and demonstration of their use in larger systems. The most obvious hole in the new design equations and topologies is the lack of measurement results. To further validate and refine the design methodology presented in this thesis a full tape-out would be required. While post-layout and corner simulations are a strong indicator of performance, a full tape out could indicate whether the transistor sizes need to be adjusted further. Measurement results for the new topologies would also be useful but the degradation in performance of the conventional latches and new topologies is expected to be similar thus not necessarily required. With this said, a further analysis of how the addition of an extra switch affects the power consumed in driving the clocks would be useful. However, this is likely better done at the system level, which leads to the most important future work.

Utilizing the new design methodology and topologies in larger systems, such as a wireline transceiver, is the next step. The improvements made in CS latch design are a good step, but ultimately they will be most useful if they can help push a SerDes to 56Gb/s or 112Gb/s at a power efficiency of 1mW/Gb/s. When utilized in a system, it must be understood where each CS latch is best used based on factors such as required output swing and impedance of the preceding and following stage. Further analysis of the CS latches when loaded (as in a system) would, therefore, be appropriate. The design equations should account for the load and common-mode changes. Furthermore, the latches have been simulated for common-mode voltages ranging from 500mV-850mV, which should be an appropriate range for a system. Nonetheless, further analysis of the latches as part of a system would be useful.

The insight learned from designing the low-power CS latches and analyzing the transient operation of the high-swing latches should be sufficient in helping a designer size high-swing latches. However, developing an in depth design methodology for the high-swing latches could also prove useful. Design equations for the new CS latch with an output switch could also be developed. This would help a designer determine what size of output switch is optimal for reducing M_3/M_4 as much as possible. Though initial results are very promising, design equations could help maximize the potential

of this new topology.

There are several other potential issues with CS latches that could be explored more in depth. Power supply noise, as it pertains to CS circuits, has not been addressed in this thesis. Given the differential nature of CS latches, it is not expected to significantly affect their performance. Nonetheless, further exploration of power supply noise in relation to CS circuits could be done in the future. Sampling aperture has also not been addressed in this thesis. The reset state of CS latches should ensure that all switches see the same voltage to start each evaluation period. As such, sampling aperture is not expected to be an issue. However, this could be explored more in depth in the future as well. Kickback noise has not been addressed in depth in this thesis. Kickback noise will certainly need to be measured in the future, but this will best be done when CS latches are used within a system. One known drawback of CS latches is the requirement for full-swing clocks. While this requirement cannot be eliminated, system improvements can certainly alleviate this issue.

More work needs to be done to understand how these CS latches scale to future technologies, such as 7nm FinFET. Some designs have expressed trouble using CML in FinFET [28], so an alternative approach may be needed. At first look, CS latches translate to FinFET nicely. Given the reset state of CS latches and the fact that not all transistors are operating at the same time, there are effectively only two stacked transistors during evaluation. This suggests CS lends itself nicely to FinFET and lower power supplies can be tolerated.

List of References

- [1] *200GbE (200 Gbit/s) over single-mode fiber and 400GbE (400 Gbit/s) over optical physical media*, IEEE Std. 802.3bs, 2017.
- [2] “Ethernet Roadmap,” Ethernet Alliance, 2016.
- [3] H. Zhang *et al.*, “PAM4 Signaling for 56G Serial Link Applications,” in *Design Conf.*, Jan. 2016.
- [4] P. Chiang *et al.*, “60Gb/s NRZ and PAM4 Transmitters for 400GbE in 65nm CMOS,” in *Intl. Solid-State Circuits Conf.*, pp. 42–44, Feb. 2014.
- [5] J. Lee *et al.*, “Design of 56 Gb/s NRZ and PAM4 SerDes Transceivers in CMOS Technologies,” in *IEEE J. of Solid-State Circuits*, vol. 50, no. 9, pp. 2061–2073, Sept. 2015.
- [6] Y. Frans *et al.*, “A 56 Gb/s PAM4 Wireline Transceiver Using a 32-Way Time-Interleaved SAR ADC in 16-nm FinFET,” in *IEEE J. of Solid-State Circuits*, vol. 52, no. 4, pp. 1101–1110, Apr. 2017.
- [7] J. Im *et al.*, “A 40-to-56 Gb/s PAM-4 Receiver with Ten-Tap Direct Decision-Feedback Equalization in 16-nm FinFET,” in *IEEE J. of Solid-State Circuits*, vol. 52, no. 12, pp. 3486–2502, Dec. 2017.
- [8] F. O’Mahony, “Wireline - 2017 Trends,” in *Intl. Solid-State Circuits Conf. Trends 2017*, pp. 16–18, Oct. 2016.
- [9] K. Fukuda *et al.*, “A 12.3-mW 12.5-Gb/s Complete Transceiver in 65-nm CMOS Process,” in *IEEE J. of Solid-State Circuits*, vol. 45, no. 12, pp. 2838–2849, Dec. 2010.
- [10] K. Kaviani *et al.*, “A 0.4-mW/Gb/s Near-Ground Receiver Front-End with Replica Transconductance Termination Calibration for a 16-Gb/s Source-Series Terminated Transceiver,” in *IEEE J. of Solid-State Circuits*, vol. 48, no. 3, pp. 636–648, Mar. 2013.
- [11] J. Polution *et al.*, “A 0.54 pJ/b 20 Gb/s Ground-Referenced Single-Ended Short-Reach Serial Link in 28 nm CMOS for Advanced Packaging Applications,” in *IEEE J. of Solid-State Circuits*, vol. 48, no. 12, pp. 3206–3218, Dec. 2013.
- [12] B. Dehlaghi and A. Carusone, “A 20 Gb/s 0.3 pJ/b Single-Ended Die-to-Die Transceiver in 28 nm-SOI CMOS,” in *Custom Integr. Circuits Conf.*, pp. 1–4, Sept. 2015.

- [13] J. W. Jung and B. Razavi, "A 25-Gb/s 5-mW CMOS CDR/Deserializer," in *IEEE J. of Solid-State Circuits*, vol. 48, no. 3, pp. 684–697, Mar. 2013.
- [14] J. W. Jung and B. Razavi, "A 25 Gb/s 5.8 mW CMOS Equalizer," in *IEEE J. of Solid-State Circuits*, vol. 50, no. 2, pp. 515–526, Feb. 2015.
- [15] A. Manian and B. Razavi, "A 40Gb/s 14mW CMOS Wireline Receiver," in *Intl. Solid-State Circuits Conf.*, pp. 412–414, Feb. 2016.
- [16] A. Manian and B. Razavi, "A 40-Gb/s 14-mW CMOS Wireline Receiver," in *IEEE J. of Solid-State Circuits*, vol. 52, no. 9, pp. 2407–2421, Sept. 2017.
- [17] A. Emami *et al.*, "A 6.0-mW 10.0-Gb/s Receiver With Switched-Capacitor Summation DFE," in *IEEE J. of Solid-State Circuits*, pp. 889–896, Mar. 2007.
- [18] T. Toifl *et al.*, "A 2.6 mW/Gbps 12.5 Gbps RX With 8-Tap Switched-Capacitor DFE in 32 nm CMOS," in *IEEE J. of Solid-State Circuits*, pp. 897–910, Mar. 2012.
- [19] J. W. Jung, "A 25-Gb/s 5-mW CDR/Deserializer in 65-nm Technology," Ph.D. dissertation, Electr. Eng. Dept., UCLA, Los Angeles, California, 2012.
- [20] A. Manian, "Low-Power Techniques for CMOS Wireline Receivers," Ph.D. dissertation, Electr. Eng. Dept., UCLA, Los Angeles, California, 2016.
- [21] J. Savoj and B. Razavi, "A 10-Gb/s CMOS clock and data recovery circuit with a half-rate linear phase detector," in *IEEE J. of Solid-State Circuits*, pp. 761–768, May 2001.
- [22] B. Razavi, "Charge Steering: A Low-Power Design Paradigm," in *Custom Integr. Circuits Conf.*, pp. 1–8, Sept. 2013.
- [23] T. Kobayashi *et al.*, "A Current-Controlled Latch Sense Amplifier and a Static Power-Saving Input Buffer for Low-Power Architecture," in *IEEE J. of Solid-State Circuits*, pp. 523–527, Apr. 1993.
- [24] A. Yusuf and Y. Leblebici, *Extreme Low-Power Mixed Signal IC Design*. New York, New York: Springer, 2010.
- [25] S. Terry, "An Introduction to the EKV Model and a Comparison of EKV to BSIM," University of Tennessee, 2005.
- [26] J. Montanaro *et al.*, "A 160-MHz 32-b 0.5-W CMOS RISC microprocessor," in *IEEE J. of Solid-State Circuits*, pp. 1703–1714, Nov. 1996.
- [27] M. Erett *et al.*, "A 126mW 56Gb/s NRZ Wireline Transceiver for Synchronous Short-Reach Applications in 16nm FinFET," in *Intl. Solid-State Circuits Conf.*, pp. 274–276, Feb. 2018.
- [28] S. Chen *et al.*, "A 4-to-16GHz Inverter-Based Injection-Locked Quadrature Clock Generator with Phase Interpolators for Multi-Standard I/Os in 7nm FinFET," in *Intl. Solid-State Circuits Conf.*, pp. 390–392, Feb. 2018.

- [29] A. Abidi and H. Xu, “Understanding the Regenerative Comparator Circuit,” in *Custom Integr. Circuits Conf.*, pp. 1–8, Sept. 2014.
- [30] B. Razavi, *Design of Analog CMOS Integrated Circuits*. New York, New York: McGraw-Hill Education, 2017.